Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze e tecnologie agrarie, ambientali e alimentari

Ciclo XXVII

Settore Concorsuale di afferenza: 07/B2
Settore Scientifico disciplinare:AGR/05

Modelling biogeochemical cycles in forest ecosystems: a Bayesian approach

Presentata da: Dott. Maurizio Bagnara

<table>
<tr><td>Coordinatore Dottorato</td><td>Relatore</td></tr>
<tr><td>Chiar.mo Prof. Giovanni Dinelli</td><td>Chiar.mo Prof. Federico Magnani</td></tr>
<tr><td></td><td>Correlatori</td></tr>
<tr><td></td><td>Dott. Matteo Sottocornola
Dott. Damiano Gianelle</td></tr>
</table>

Esame finale anno 2015

"Remember that all models are wrong;

the practical question is how wrong do they have to be

to not be useful. "

(George E.P. Box)

Table of contents

Chapter 1: Introduction

1. The importance of forest ecosystems

Forest ecosystems are an environmental and economic resource which is spread on a local to global scale. Some authors suggest their extension is about 40% of the Earth's ice-free land surface but it is less than the potential extension, due to human disturbances (Waring and Running, 1998). They are important as an economic resource, prevent soil erosion, contribute to maintain biodiversity and play a social role as recreational areas. In addition, they also play a major role in both the water and carbon cycles as they sequester large amounts of carbon (Dixon et al., 1994) and have a strong impact on evapotranspiration and runoff (Huntington, 2006; Vörösmarty and Sahagian, 2000). In the global carbon (C) budget, the terrestrial biosphere is thought to be a significant C sink, sequestering about 2.8 PgC yr^{-1} (Canadell et al., 2007; Field, 2001; Myneni et al., 2001). Although the ecosystems responsible of this carbon sequestration is still debated, it has been suggested that temperate and boreal ecosystems of the northern hemisphere contribute with an uptake of approximately 0.6–0.7 PgC yr^{-1} (Goodale et al., 2002). Given that this estimates are still largely uncertain, many ecological studies aim at increasing our understanding of ecosystem functioning and at quantifying ecosystem processes.

There are several methods to assess the amount of C sequestered or released by forest ecosystems, including but not limited to chamber measurements, FACE experiments, dendro-ecological studies, Eddy-covariance systems, proximal and remote sensing techniques, and modelling tools. In this thesis, the focus will be on the latter.

2. Forest modelling

Forest models are tools for explaining and predicting the dynamics of forest ecosystems. They simulate forest behavior by integrating information on the underlying processes in trees, soil and atmosphere. Their complexity is a strength, allowing them, in principle, to reproduce the complex dynamics of forest ecosystems in detail, but it is also a weakness because it makes their use and evaluation difficult (van Oijen et al., 2005). The advantage of using models is that they can be

extended across large spatial domains and into the future, given the relevant driving variables (Rastetter et al., 2003; Running et al., 1999; White et al., 2000): forecasts are possible because models incorporate a representation of the simulated system and its dynamics. Another important advantage of modelling is the possibility of obtaining insights about ecosystems behavior and their responses to environmental variables (Grant et al., 2012; Grassi and Magnani, 2005; Magnani et al., 2002; Oulehle et al., 2011), as well as the quantification of the importance of these variables through the application of sensitivity analysis techniques (Davi et al., 2006; Dufrêne et al., 2005; Fu et al., 2012; Wu and Liu, 2012). Their disadvantage is that the process of model construction is arguably subjective: Occam's razor – making models as simple as possible, but not simpler – is a useful guiding principle, but there is always a danger that the model's representation of the system is not accurate. Other problems include model parameterization: generally, model parameters are unknown and have to be derived from data, so there is always a danger that poorly defined parameters will be 'tuned' to give good output. Models are generally parameterized with some subset of observational data, and tested against remaining data, and such tests are designed to show that the model can effectively describe the observed system by demonstrating a strong correlation, or a low mean error, between prediction and observation. When several parameters are tuned, the right answer may be generated for the wrong reason (Williams et al., 2001).

Most published models of forest growth are concerned with tree physiology or nutrient cycling, concentrating, respectively, on photosynthetic carbon gain and allocation or on decomposition and nutrient uptake processes (Nightingale et al., 2004). Several reviews on forest model have been written, trying to define categories of models. For example, Nightingale et al. (2004) divided the existing models in 4 "organizational levels":

1. Leaf–tree organizational level
2. Plot/stand organizational level
3. Regional organizational level
4. Ecosystem organizational level

Landsberg (2003) divided the existing models depending on the type of user groups at which they refer:

1. Models relating to industry

2. Models relating to the broader public community

3. Models relating to the academic world

while Pretzsch et al. (2008) gave an "historical" division, based on model complexity and aims:

1. Maps and yield tables

2. Growth- and yield simulators

3. Individual tree-orientated management models

4. Gap and hybrid models

5. Matter-balance models

6. Landscape models

7. Visualization models

where the categories from 3 to 6 comprises process-based models with increasing complexity. Existing process-based models were also reviewed by Mäkelä et al. (2000), particularly focusing on their application to forest management.

Studies of particular interest for modellers are those focusing on model comparison (several models tested on the same site) and on model evaluation (one model tested on several sites), as they are useful to know how models work and which are their main strengths and weaknesses. An interesting study of the first type was carried out by Kramer et al. (2002), who compared the performances in predicting CO_2 and H_2O fluxes of 6 process-based models at 6 European forest sites, on the basis of 3 criteria: generality, accuracy, and realism. An example of the second type of study was performed by Landsberg et al. (2003), testing the performances of the model 3-PG in describing and predicting forest growth at several sites (ranging from sub-tropical Africa and Australia to northern Europe).

The classification made by Nightingale et al. (2004) is particularly useful to define which kind of model is most suitable for use depending on the temporal and spatial scale at which researchers are working: a brief description of the 4 categories presented above follows.

1. *Leaf–tree organizational level*

   Models at this organizational level have been designed to:

   1) represent the effects of atmospheric carbon, irradiance, temperature, nutrients and water deficits on both the light and dark reactions of carbon fixation at the leaf level;

   2) scale photosynthetic rates from the leaf to canopy level;

   3) simulate the establishment, annual diameter growth and mortality of individual trees;

   4) calculate radiation, absorption, net photosynthesis and transpiration rates in individual tree canopies as well as the allocation of dry matter to the component parts of the tree and hence the growth patterns of trees;

   5) examine the mechanisms by which plants regulate their carbon, water and nutrient cycles under different seasonal climatic conditions to mitigate damage caused by pollutants;

   6) examine changes in the fluxes and allocation of carbon and nitrogen among foliage, fine roots, stems and soils.

   Models at this organizational level include ECOSYS, WIMOVAC, TREGROW, MAESTRO, BEX, BIOMASS, PIPESTEM, GEM, FVS, DAYTRANS/PSN and ZELIG.

2. *Plot/stand organizational level*

   The majority of process-based forest growth, or ecosystem, models simulate the growth of stands or forest plots at spatial scales of metres to 1 km^2. Models at this organizational level have been designed to:

   1) examine the effects on stand growth due to weather variables, soil water, air pollutants, rooting characteristics and nutrient uptake;

   2) calculate the carbon balance of the stand based on photosynthesis and respiration

3) simulate community dynamics, growth parameters, biomass, photosynthesis and productivity of forest stands and calculate the hydrological balance and growth of a forest;

4) estimate the spatial and temporal distribution of major carbon and nitrogen fluxes and pool sizes in forest ecosystems.

Models within this organizational level include HYDRALL, SPA, DAYTRANS/PSN, FORGRO, FORMIND, JABOWA, FORMIX, HYBRID, FOREST-BGC, 3-PG, 3-PGS, Prelued, CASTANEA, CENTURY, BEPS and FORECAST. Several ecosystem models operate at the scale between plot/stand and regional levels and include FBM, NASA CASA, PnET, CASA, TEM, CARAIB, BIOME-BGC and FORESEE.

3. *Regional organizational level*

Models that operate specifically at the regional scale (1–10 km^2) are similar to those that function at the plot level. Models at this organizational level have essentially been designed to

1) simulate seasonal patterns in live biomass, annual plant production and soil carbon and nitrogen levels;

2) simulate global scale patterns of NPP and sensitivity of global ecosystems to changes in climate.

Examples of models at the regional organizational level are SiB, DEMETER G'DAY, LPJ-GUESS and more recently CENTURY.

4. *Ecosystem organizational level*

Models that operate at this organizational level may also be applied to all global biomes. This organizational modelling level has essentially been designed to:

1) represent global NPP, and examine the effects of different anthropogenic carbon dioxide emissions and mitigation scenarios on atmospheric CO^2

2) provide long-term forecasts of global changes, as well as generalized estimates of global primary productivity and vegetation and soil carbon pools and fluxes for each biome represented in the model output.

Examples of global scale productivity models are CoupModel, ORCHIDEE, GLOPEM and GLOCO.

Many of the models of forest growth and biogeochemical cycles cited above, developed in the last 30 years, are complex research tools that replicate forest physiological processes. These are typically detailed, multi-variable models that need large datasets of environmental drivers and careful species-specific parameterization (Landsberg and Waring, 1997). Therefore, to make them suitable for a wider range of purposes and accessible to a wider range of users, a process of simplification started in the 1990s (Landsberg and Waring, 1997; White and Running, 1994) with the aim of developing generalized models that could be of use in applied forest management. There are advantages in using simpler models to estimate ecosystem carbon state variables: 1) there is a trade-off between model complexity, such as the number of model parameters, and a model's ability to reproduce observations (e.g. Akaike, 1974), therefore a low-complexity model is preferable when it can reproduce ecosystem observations with comparable skill; 2) complex models are often computationally expensive, and this is an inhibiting factor when using iterative methods to estimate model parameters and their uncertainty (van Oijen et al., 2005). One step in this direction was represented by the creation of hybrid models like FORCYTE-11 (Kimmins, 1986), that combine the predictive power of process-based models with the short-term believability of mensuration-based models (Kimmins et al., 1999; Landsberg, 2003). Unlike full process-based models, hybrid models are based on the principle that only the processes that are expected to change would be included in the modelling effort (Kimmins et al., 2008). The effort towards simplification is not limited to hybrid models: a widely used group of simple models for the prediction of GPP is based on the concept of light use efficiency (LUE). These models assume that vegetation has a potential LUE, which departs from the actual LUE because the latter is affected by differences in intercepted photosynthetically active radiation (PAR) and environmental constraints (Kumar and Monteith, 1981; Landsberg and Waring, 1997; Monteith and Moss, 1977). Therefore, the optimal LUE is decreased by modifying factors that account for sub-optimal conditions for photosynthesis

(Landsberg and Waring, 1997; McMurtrie et al., 1994). LUE models mainly rely on a simplified representation of physiological processes based on empirical parameters and their mathematical structure is often quasi- or totally multiplicative. As a consequence, LUE models typically require limited input data and are computationally efficient. Some examples of these models are 3PG (Landsberg and Waring, 1997), EC–LUE (Yuan et al., 2007), C-Fix (Veroustraete et al., 1994), CFLUX (Turner et al., 2006) and Prelued (Mäkelä et al., 2008).

3. The Bayesian framework

Many different forest models have been published, but in most cases we know little about their reliability. Van Oijen et al. (2004) identified a range of biogeochemical processes, already incorporated in models, for which data are lacking: our incomplete knowledge constitutes uncertainty that we may represent in the form of probability distributions. Our uncertainties are linked: uncertain parameters imply uncertain predictions and uncertainty about the real world implies uncertainty about model structure and parameterization. Because of these linkages, our methods for model parameterization, uncertainty analysis, sensitivity analysis, prediction, testing and comparison with other models need to be based on a consistent quantification of uncertainty. Bayesian statistics aims to provide such "rules and procedures for disciplined uncertainty accounting" (Bernardo and Smith, 2009). It is a method, applicable to all models, that quantifies output uncertainty, identifies key parameters and variables, and is of use in model improvement and model selection. In the last decade, the Bayesian approach to model calibration has been widely applied to fulfill that purpose (amongst others, Bloom and Williams, 2015; Milne et al., 2014; van Oijen et al., 2011, 2005).

Bayesian calibration is the application of probability theory to parameter estimation (Jaynes, 2003; Sivia, 1996), and this method finds increasing use in ecological modelling (Ogle and Barber, 2008; Ogle, 2009; van Oijen et al., 2005). The main characteristic of the Bayesian approach is that it quantifies model inputs and outputs in the form of probability distributions, and applies the rules of probability theory to update the distributions when new data are obtained (Sivia, 1996).

7

Uncertainty about parameters is represented as a joint probability distribution for the possible parameter values. Bayes' Theorem is used to determine how this distribution changes in the light of new data:

$$P(\theta|D) \propto P(\theta)\, P(D|\theta)$$

where $P(\theta)$ and $P(\theta|D)$ are the prior and posterior distributions for the parameters θ, i.e. before and after conditioning on the data D. The factor that modifies the prior, $P(D|\theta)$, is the likelihood function, which is the probability of the data D given the model output generated by the parameter vector θ, and it accounts for possible measurement error. A formal likelihood function, integrating to unity in data space, needs to be used to be consistent with the probability calculus, allowing Bayes' Theorem to be applied. The prior probability distribution for the parameters of a model, $P(\theta)$, reflects the modeller's uncertainty about parameter values before using the data.

Bayesian approaches have been much less frequently used in forest research (Ghazoul and McAllister, 2003) than in the other environmental sciences (Ellison, 2004, 1996; Reckhow, 1990), because treating everything probabilistically, even if the model has many different inputs and outputs, makes Bayesian calibration computationally demanding. This has hampered its practical use in the past, and a very few examples exist for forest models with many parameters (Svensson et al., 2008). However, recent developments in sampling-based evaluation of probability distributions, in particular Markov Chain Monte Carlo techniques, can help alleviate the computational problem (van Oijen et al., 2005).

The procedure for Bayesian calibration begins with quantifying the uncertainty about parameter values in the form of so-called prior probability distributions. Then, the measured data is used on the output variables to calibrate the model. The calibration, which is done by application of Bayes' Theorem, yields an updated posterior distribution for the parameters. This focus on quantifying uncertainty in the form of probability distributions distinguishes Bayesian calibration from parameter estimation methods that focus on goodness-of-fit, and the predictive uncertainty of the model can be determined by running the model with different parameter settings sampled from the

posterior parameter distribution. The application of a probabilistic framework affects not only the way we model, but also the way we report our model and modeling results. When communicating model results, we need to present more than just the most likely predictions, but uncertainty intervals should also be generated together with the full joint probability distribution of the parameters, as produced by the Bayesian calibration (van Oijen et al., 2005). New model users can use this parameter probability distribution as their prior, and update it by applying Bayes' Theorem to their own measured output data, as long as it is made clear which data have already been used, to avoid double-counting.

Implementations of Bayesian calibration rely mainly on Markov Chain Monte Carlo (MCMC) algorithms for sampling the parameter space to obtain posterior probability distributions for the model parameters (Hill et al., 2012; Rahn et al., 2011; Sacks et al., 2006). A Markov chain is a random process that undergoes transitions from one state to another on a state space (the set of values which a process can take): in a Markov chain, the probability distribution of the next state depends only on the current state and not on the sequence of states that preceded it . This specific property is called the "Markov property" (Bryson, 1975). Markov chains have many applications as statistical models of real-world processes, and are more and more applied to different kind of fields and topics: among others, physic, chemistry, statistics, biology and genetics (Asmussen and Glynn, 2007). Markov Chain Monte Carlo (MCMC) methods are a class of algorithms for sampling from a prior probability distribution towards a desired posterior probability distribution (known as the equilibrium distribution) based on constructing a Markov chain. MCMC methods allow us to represent any posterior distribution whatever its shape, improving on older methods that only subdivide the parameter space into accepted and rejected regions (Mäkelä, 1988). They are primarily used for calculating numerical approximations of multi-dimensional integrals, for example in Bayesian statistics, computational physics, computational biology and computational linguistics (Andrieu et al., 2003). When an MCMC method is used for approximating a multi-dimensional integral (which is the case for complex models with high number of parameters),

an ensemble of "walkers" move around randomly. At each point where a walker steps, the integrand value at that point is counted towards the integral. The walker then may make a number of tentative steps around the area, looking for a place with a reasonably high contribution to the integral to move into next. The state of the chain after a number of steps is then used as a sample of the desired posterior distribution, and the quality of the sample improves as a function of the number of steps (Gilks et al., 1996). The more difficult problem when constructing a Markov chain for sampling from a desired posterior distribution is to determine how many steps are needed to converge to that distribution within an acceptable error: typically, MCMC sampling can only approximate the target distribution, Moreover, the efficiency of the MCMC technique is also highly dependent on the model structure: the high correlations between parameters that are induced by a multiplicative model structure generally make the convergence of the MCMC more difficult (Browne et al., 2009; Gilks and Roberts, 1996). This is one of the reasons why the application of the Bayesian method to LUE-based models (often multiplicative) is not as common as its application to process-based models, with a very few studies heading in this direction (Minunno et al., 2010; Still et al., 2004; Xenakis et al., 2008). Different methods have been implemented to avoid or reduce such problems: the use of very long chains (Gilks et al., 1996), model re-parameterization to avoid strong correlations (Buzzi-Ferraris and Manenti, 2010; Gilks et al., 1996), and the use of more efficient algorithms (Gilks et al., 1996; ter Braak, 2006).

Even if this thesis focuses on the application of the MCMC method, the Bayesian calibration in general does not necessarily rely on it. Two other methods are sometimes applied to calibrate ecosystem model: the adjoint method (Zhu et al., 2014) and the Kalman filter (Gao et al., 2011). These data assimilation techniques mentioned above are special cases of Bayesian calibration (Wikle and Berliner, 2007), where a prior probability distribution for parameters is specified and updated using Bayes' Theorem. However, in contrast to the MCMC approach, the old data assimilation methods – though computationally efficient - require assumptions of linearity and Gaussian distributions that are restrictive and inappropriate in the case of highly nonlinear models

like many forest models. Therefore such methods are common in state estimation of computationally demanding models such as General Circulation Models, but they are not common in parameter estimation of ecosystem models. The MCMC method allows for any type of prior and posterior distribution, including asymmetric and multimodal ones, in contrast to the other methods. Moreover, the sample from the posterior distribution generated by MCMC represents the full posterior probability distribution, in contrast to the adjoint method which only provides an estimate of the mode, and model uncertainties can only be fully assessed with such global methods instead of local ones.

4. Objectives

This study aims at testing the Bayesian procedure for calibration to different types of forest models, to evaluate their performances and the uncertainties associated with them. In particular, the main objectives were the following:

- to apply a Bayesian framework to calibrate forest models and test their performances in different biomes and different environmental conditions;

- to identify and solve structure-related issues in simple models;

- to identify the advantages of additional information made available when calibrating forest models with a Bayesian approach.

5. Thesis outline

This thesis is organized in 5 Chapters which summarize the work carried out in three years of activity. After this first introductory section, Chapter 2 reports a case study on the application of the LUE-based model Prelued (Mäkelä et al., 2008) on several Italian Eddy-covariance forest sites, covering a wide range of biomes and climatic and environmental conditions (Bagnara et al., 2014a). Chapter 3 focuses on the structure-related issues of the LUE-based model Prelued when calibrated with a Bayesian approach and on the procedures developed to solve them (Bagnara et al., 2014b). Chapter 4 reports a case study on the application of the process-based model HYDRALL (Magnani

et al., 2002) on an alpine forest site, calibrated on multiple variables with a Bayesian approach.

General conclusions are summarized in Chapter 5.

6. References

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Trans. Automat. Contr. 19, 716–723.

Andrieu, C., de Freitas, N., Doucet, A., Jordan, M., 2003. An Introduction to MCMC for Machine Learning. Mach. Learn. 50, 5–43. doi:10.1023/A:1020281327116

Asmussen, S., Glynn, P.W., 2007. Stochastic simulation: Algorithms and analysis. Springer Science & Business Media.

Bagnara, M., Sottocornola, M., Cescatti, A., Minerbi, S., Montagnani, L., Gianelle, D., Magnani, F., 2014a. Bayesian optimization of a light use efficiency model for the estimation of daily gross primary productivity in a range of Italian forest ecosystems. Ecol. Modell. doi:10.1016/j.ecolmodel.2014.09.021

Bagnara, M., Van Oijen, M., Cameron, D., Gianelle, D., Magnani, F., Sottocornola, M., 2014b. A user-friendly forest model with a multiplicative mathematical structure: a Bayesian approach to calibration. Geosci. Model Dev. Discuss. 7, 6997–7031. doi:10.5194/gmdd-7-6997-2014

Bernardo, J.M., Smith, A.F.M., 2009. Bayesian theory. John Wiley & Sons.

Bloom, A.A., Williams, M., 2015. Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological" common sense" in a model–data fusion framework. Biogeosciences 12, 1299–1315.

Browne, W.J., Steele, F., Golalizadeh, M., Green, M.J., 2009. The use of simple reparameterizations to improve the efficiency of Markov chain Monte Carlo estimation for multilevel models with applications to discrete time survival models. J. R. Stat. Soc. Ser. A (Statistics Soc. 172, 579–598.

Bryson, A.E., 1975. Applied optimal control: optimization, estimation and control. CRC Press.

Buzzi-Ferraris, G., Manenti, F., 2010. Better reformulation of kinetic models. Comput. Chem. Eng. 34, 1904–1906.

Canadell, J.G., Le Quéré, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A., Marland, G., 2007. Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. Proc. Natl. Acad. Sci. 104 , 18866–18870. doi:10.1073/pnas.0702737104

Davi, H., Dufrene, E., Francois, C., Lemaire, G., Loustau, D., Bosc, a, Rambal, S., Granier, a, Moors, E., 2006. Sensitivity of water and carbon fluxes to climate changes from 1960 to 2100 in European forest ecosystems. Agric. For. Meteorol. 141, 35–56.

Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C., Wisniewski, J., 1994. Carbon pools and flux of global forest ecosystems. Science (80-.). 263, 185–190.

Dufrêne, E., Davi, H., François, C., Maire, G. Le, Dantec, V. Le, Granier, a., 2005. Modelling carbon and water cycles in a beech forestPart I: Model description and uncertainty analysis on modelled NEE. Ecol. Modell. 185, 407–436.

Ellison, A.M., 2004. Bayesian inference in ecology. Ecol. Lett. 7, 509–520.

Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecol. Appl. 6, 1036–1046.

Field, C.B., 2001. Plant Physiology of the "Missing" Carbon Sink. Plant Physiol. 125, 25–28.

Fu, Y.H., Campioli, M., Van Oijen, M., Deckmyn, G., Janssens, I.A., 2012. Bayesian comparison of six different temperature-based budburst models for four temperate tree species. Ecol. Modell. 230, 92–100.

Gao, C., Wang, H., Weng, E., Lakshmivarahan, S., Zhang, Y., Luo, Y., 2011. Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics. Ecol. Appl. 21, 1461–1473.

Ghazoul, J., McAllister, M., 2003. Communicating complexity and uncertainty in decision-making contexts: Bayesian approaches to forest research. Int. For. Rev. 5, 9–19.

Gilks, W., Roberts, G., 1996. Strategies for Improving MCMC., in: Gilks, W., Richardson, S., Spiegelhalter, D. (Eds.), Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, FL., pp. 89–114.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov chain Monte Carlo in practice. CRC press.

Goodale, C.L., Apps, M.J., Birdsey, R.A., Field, C.B., Heath, L.S., Houghton, R.A., Jenkins, J.C., Kohlmaier, G.H., Kurz, W., Liu, S., Nabuurs, G.-J., Nilsson, S., Shvidenko, A.Z., 2002. Forest Carbon Sinks In The Northern Hemisphere. Ecol. Appl. 12, 891–899. doi:10.1890/1051-0761(2002)012[0891:FCSITN]2.0.CO;2

Grant, R.F., Baldocchi, D.D., Ma, S., 2012. Ecological controls on net ecosystem productivity of a seasonally dry annual grassland under current and future climates: Modelling with ecosys. Agric. For. Meteorol. 152, 189–200.

Grassi, G., Magnani, F., 2005. Stomatal, mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees. Plant, Cell Environ. 28, 834–849. doi:10.1111/j.1365-3040.2005.01333.x

Hill, T.C., Ryan, E., Williams, M., 2012. The use of CO2 flux time series for parameter and carbon stock estimation in carbon cycle research. Glob. Chang. Biol. 18, 179–193.

Huntington, T.G., 2006. Evidence for intensification of the global water cycle: Review and synthesis. J. Hydrol. 319, 83–95. doi:http://dx.doi.org/10.1016/j.jhydrol.2005.07.003

Jaynes, E.T., 2003. Probability theory: the logic of science. Cambridge university press.

Kimmins, J.P., 1986. Predicting the consequences of intensive forest harvesting on long-term productivity: the need for a hybrid model such as FORCYTE-11. Rapp. Institutionen Ekol. och Miljovard, Sveriges Lantbruksuniversitet 31–84.

Kimmins, J.P., Blanco, J.A., Seely, B., Welham, C., Scoullar, K., 2008. Complexity in modelling forest ecosystems: How much is enough? For. Ecol. Manage. 256, 1646–1658. doi:10.1016/j.foreco.2008.03.011

Kimmins, J.P., Mailly, D., Seely, B., 1999. Modelling forest ecosystem net primary production: the hybrid simulation approach used in forecast. Ecol. Modell. 122, 195–224. doi:10.1016/S0304-3800(99)00138-6

Kramer, K., Leinonen, I., Bartelink, H.H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, a. J., Froer, O., Gracia, C. a., Granier, a., Grunwald, T., Hari, P., Jans, W., Kellomaki, S., Loustau, D., Magnani, F., Markkanen, T., Matteucci, G., Mohren, G.M.J., Moors, E., Nissinen, a., Peltola, H., Sabate, S., Sanchez, a., Sontag, M., Valentini, R., Vesala, T., 2002. Evaluation of six process-based forest growth models using eddy-covariance measurements of CO_2 and H_2O fluxes at six forest sites in Europe. Glob. Chang. Biol. 8, 213–230. doi:10.1046/j.1365-2486.2002.00471.x

Kumar, M., Monteith, J.L., 1981. Remote sensing of crop growth, in: Plants and the Daylight Spectrum. Academic Press: London, pp. 133–144.

Landsberg, J., 2003. Modelling forest ecosystems: state of the art, challenges, and future directions. Can. J. For. Res. 33, 385–397. doi:10.1139/x02-129

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

Landsberg, J., Waring, R.H., Coops, N.C., 2003. Performance of the forest productivity model 3-PG applied to a wide range of forest types. For. Ecol. Manage. 172, 199–214. doi:10.1016/S0378-1127(01)00804-0

Magnani, F., Grace, J., Borghetti, M., 2002. Adjustment of tree structure in response to the environment under hydraulic constraints. Funct. Ecol. 16, 385–393. doi:10.1046/j.1365-2435.2002.00630.x

Mäkelä, A., 1988. Performance analysis of a process-based stand growth model using Monte Carlo techniques. Scand. J. For. Res. 3, 315–331.

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen, P., 2000. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20, 289–298.

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. Chang. Biol. 14, 92–108. doi:10.1111/j.1365-2486.2007.01463.x

McMurtrie, R.E., Gholz, H.L., Linder, S., Gower, S.T., 1994. Climatic factors controlling the productivity of pine stands : a model-based analysis. Ecol. Bull. 43, 173–188.

Milne, A.E., Glendining, M.J., Bellamy, P., Misselbrook, T., Gilhespy, S., Casado, M.R., Hulin, A., Van Oijen, M., Whitmore, A.P., 2014. Analysis of uncertainties in the estimates of nitrous oxide and methane emissions in the UK's greenhouse gas inventory for agriculture. Atmos. Environ. 82, 94–105.

Minunno, F., Xenakis, G., Perks, M.P., Mencuccini, M., 2010. Calibration and validation of a simplified process- based model for the prediction of the carbon balance of Scottish Sitka spruce (Picea sitchensis) plantations. Can. J. For. Res. 40, 2411–2426. doi:10.1139/X10-181

Monteith, J.L., Moss, C.J., 1977. Climate and the Efficiency of Crop Production in Britain [and Discussion]. Philos. Trans. R. Soc. Lond. B 281, 277–294. doi:10.1098/rstb.1977.0140

Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., Liski, J., Zhou, L., Alexeyev, V., Hughes, M.K., 2001. A large carbon sink in the woody biomass of Northern forests. Proc. Natl. Acad. Sci. 98 , 14784–14789. doi:10.1073/pnas.261555198

Nightingale, J.M., Phinn, S.R., Held, A.A., 2004. Ecosystem process models at multiple scales for mapping tropical forest productivity. Prog. Phys. Geogr. 28, 241–281. doi:10.1191/0309133304pp411ra

Ogle, K., 2009. Hierarchical Bayesian statistics: merging experimental and modeling approaches in ecology. Ecol. Appl. 19, 577–581.

Ogle, K., Barber, J.J., 2008. Bayesian data—model integration in plant physiological and ecosystem ecology, in: Progress in Botany. Springer, pp. 281–311.

Oulehle, F., Evans, C.D., Hofmeister, J., Krejci, R., Tahovska, K., Persson, T., Cudlin, P., Hruska, J., 2011. Major changes in forest carbon and nitrogen cycling caused by declining sulphur deposition. Glob. Chang. Biol. n/a–n/a. doi:10.1111/j.1365-2486.2011.02468.x

Pretzsch, H., Grote, R., Reineking, B., Rötzer, T., Seifert, S., 2008. Models for forest ecosystem management: a European perspective. Ann. Bot. 101, 1065–1087. doi:10.1093/aob/mcm246

Rahn, K., Butterbach-bahl, K., Werner, C., 2011. Selection of likelihood parameters for complex models determines the effectiveness of Bayesian calibration. Ecol. Inform. 6, 333–340. doi:10.1016/j.ecoinf.2011.08.002

Rastetter, E.B., Aber, J.D., Peters, D.P.C., Ojima, D.S., Burke, I.C., 2003. Using mechanistic models to scale ecological processes across space and time. Bioscience 53, 68–76.

Reckhow, K.H., 1990. Bayesian inference in non-replicated ecological studies. Ecology 71, 2053–2059.

Running, S.W., Baldocchi, D.D., Turner, D.P., Gower, S.T., Bakwin, P.S., Hibbard, K.A., 1999. A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sens. Environ. 70, 108–127.

Sacks, W.J., Schimel, D.S., Monson, R.K., Braswell, B.H., 2006. Model-data synthesis of diurnal and seasonal CO2 fluxes at Niwot Ridge, Colorado. Glob. Chang. Biol. 12, 240–259. doi:10.1111/j.1365-2486.2005.01059.x

Sivia, D.S., 1996. Data Analysis: A Bayesian Tutorial. Oxford University Press, Oxford.

Still, C.J., Randerson, J.T., Fung, I.Y., 2004. Large-scale plant light-use efficiency inferred from the seasonal cycle of atmospheric CO2. Glob. Chang. Biol. 10, 1240–1252. doi:10.1111/j.1365-2486.2004.00802.x

Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., Lindroth, a, 2008. Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand. Ecol. Modell. 213, 331–344. doi:10.1016/j.ecolmodel.2008.01.001

Ter Braak, C.J.F., 2006. A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. Stat. Comput. 16, 239–249. doi:10.1007/ s11222-006-8769-1

Turner, D.P., Ritts, W.D., Styles, J.M., Yang, Z., Cohen, W.B., Law, B.E., Thornton, P.E., 2006. A diagnostic carbon flux model to monitor the effects of disturbance and interannual variation in climate on regional NEP. Tellus B 58, 476–490. doi:10.1111/j.1600-0889.2006.00221.x

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. doi:10.1016/j.agrformet.2011.06.017

Van Oijen, M., Cannell, M.G.R., Levy, P.E., 2004. Modelling biogeochemical cycles in forests: state of the art and perspectives, in: Andersson, F., Birot, Y., Päivinen, R. (Eds.), Towards the Sustainable Use of Europe's Forests—forest Ecosystem and Landscape Research: Scientific Challenges and Opportunities. European For. Inst., Joensuu, Finland, pp. 157–169.

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, 915–27. doi:10.1093/treephys/25.7.915

Veroustraete, F., Patyn, J., Myneni, R.B., 1994. Forcing of a simple ecosystem model with fAPAR and climatic data to estimate regional scale photosynthetic assimilation. Veg. Model. Clim. Chang. Eff. 151–177.

Vörösmarty, C.J., Sahagian, D., 2000. Anthropogenic disturbance of the terrestrial water cycle. Bioscience 50, 753–765.

Waring, R.H., Running, S.W., 1998. Forest Ecosystems: Analysis at Multiple Scales. Academic Press, San Diego.

White, J.D., Running, S.W., 1994. Testing scale dependent assumptions in regional ecosystem simulations. J. Veg. Sci. 5, 687–702. doi:10.2307/3235883

White, M. a., Thornton, P.E., Running, S.W., Nemani, R.R., 2000. Parameterization and Sensitivity Analysis of the BIOME–BGC Terrestrial Ecosystem Model: Net Primary Production Controls. Earth Interact. 4, 1–85.

Wikle, C.K., Berliner, L.M., 2007. A Bayesian tutorial for data assimilation. Phys. D Nonlinear Phenom. 230, 1–16.

Williams, M., Law, B.E., Anthoni, P.M., H, M., 2001. Use of a simulation model and ecosystem flux data to examine carbon – water interactions in ponderosa pine. Tree Physiol. 287–298.

Wu, Y., Liu, S., 2012. Automating calibration, sensitivity and uncertainty analysis of complex models using the R package Flexible Modeling Environment (FME): SWAT as an example. Environ. Model. Softw. 31, 99–109. doi:10.1016/j.envsoft.2011.11.013

Xenakis, G., Ray, D., Mencuccini, M., 2008. Sensitivity and uncertainty analysis from a coupled 3-PG and soil organic matter decomposition model. Ecol. Modell. 219, 1–16. doi:10.1016/j.ecolmodel.2008.07.020

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L.L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A.H., Goulden, M.L., Hollinger, D.Y., Hu, Y., Law, B.E., Stoy, P.C., Vesala, T., Wofsy, S.C., 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agric. For. Meteorol. 143, 189–207. doi:10.1016/j.agrformet.2006.12.001

Zhu, G.F., Li, X., Su, Y.H., Zhang, K., Bai, Y., Ma, J.Z., Li, C.B., Hu, X.L., He, J.H., 2014. Simultaneously assimilating multivariate data sets into the two-source evapotranspiration model by Bayesian approach: application to spring maize in an arid region of northwestern China. Geosci. Model Dev. 7, 1467–1482. doi:10.5194/gmd-7-1467-2014

Chapter 2: Bayesian optimization of a Light Use Efficiency model for the estimation of daily Gross Primary Productivity in a range of Italian forest ecosystems

Abstract

In this study we applied a modified version of Prelued, a simple semi-empirical Light Use Efficiency (LUE) model, to eight eddy-covariance Italian sites. Since this model has been successfully applied mainly to coniferous forests located at northern latitudes, in our study we aimed to test its generality, by comparing Prelued's outputs in coniferous, broadleaf forests and in a Mediterranean macchia, at different climatic and environmental conditions. The model was calibrated for daily Gross Primary Production (GPP) observed over one year in each flux site and validated for another year. The model uncertainties on both GPP and model parameters were estimated, applying a Bayesian calibration based on a multiple chains Markov Chain Monte Carlo sampling.

The accuracy of the model estimates of daily GPP over the entire period of simulation differed widely depending on the site considered, with generally good model performance when applied to evergreen and broadleaf forests and poor performances in the Mediterranean macchia. The values of the modifiers accounting for the response to climatic variables suggested the soil water content to be non-limiting in temperate mountain evergreen but limiting in Mediterranean forests. Model uncertainties were always smaller than data uncertainties, with variable magnitude depending on the site considered. Both modelled GPP and uncertainties were largely dependent also on uncertainties on the data, which made their calculation a key process in this modelling exercise.

In conclusion, this semi-empirical model appears to be suitable for estimating daily and annual forest GPP in most of the considered sites, with the exception of Mediterranean macchias, and for supporting its application to a large range of ecosystems provided a site-specific calibration. The Bayesian calibration did not confer a clear advantage in terms of model performances in respect to

other methods used in previous studies, but allowed us to estimate uncertainties on both parameter values and model estimates, which were useful to analyse more in detail the ecosystem response to environmental drivers of GPP.

1. Introduction

The Gross Primary Production (GPP), being the largest carbon flux between the atmosphere and the biosphere, is among the main outputs of many forest ecosystem models (Foley et al., 1996; J. E. Horn and Schulz, 2011; Landsberg and Waring, 1997; Mäkelä et al., 2008). GPP is also being increasingly targeted by remote sensing applications as a proxy to assess global carbon fluxes (such as Net Ecosystem CO_2 Exchange, NEE) and plant light-use efficiency at large spatial scales (Still et al., 2004; Xiao et al., 2005). At the same time, the quantification of GPP is a challenge in most ecosystems because of its dependence on a variety of interlinked meteorological, environmental and biological drivers at several time scales.

Many of the models of forest growth and biogeochemical cycles developed in the last 30 years are complex research tools that replicate forest physiological processes. These are typically detailed, multi-variable models that need large datasets of environmental drivers and careful species-specific parameterisation (Landsberg and Waring, 1997). Therefore, a process of simplification started in the 90's (Landsberg and Waring, 1997; White and Running, 1994) with the aim of developing generalized models that could be of use in applied forest management.

One step in this direction was represented by the creation of hybrid models like FORCYTE-11 (Kimmins, 1986), that combine the predictive power of process-based models with the short-term believability of mensuration-based models (Kimmins et al., 1999; Landsberg, 2003). Unlike full process-based models, hybrid models are based on the principle that only the processes that are expected to change would be included in the modelling effort (Kimmins et al., 2008).

The effort towards simplification is not limited to hybrid models: a widely used group of simple models for the prediction of GPP is based on the concept of Light Use Efficiency (LUE). These

models assume that vegetation has a potential LUE, which departs from the actual LUE because the latter is affected by differences in intercepted Photosynthetically Active Radiation (PAR) and environmental constraints(Kumar and Monteith, 1981; Landsberg and Waring, 1997; Monteith and Moss, 1977). Therefore, the optimal LUE is decreased by modifying factors that account for sub-optimal conditions for photosynthesis (Landsberg and Waring, 1997; McMurtrie et al., 1994). LUE models mainly rely on a simplified representation of physiological processes based on empirical parameters and their mathematical structure is often quasi- or totally multiplicative. As a consequence, LUE models typically require limited input data and are computationally efficient. Some examples of these models are 3PG (Landsberg and Waring, 1997), EC-LUE (Yuan et al., 2007), C-Fix (Veroustraete et al., 1994),CFLUX(Turner et al., 2006)and Prelued (Mäkelä et al., 2008).

Despite relying on a multiplicative mathematical structure and on several empirical parameters, of which little is known in the literature, Prelued has been successfully applied to several ecosystems, but mainly in evergreen coniferous forest from northern latitudes(Mäkelä et al., 2008; McCallum et al., 2013; Peltoniemi et al., 2012a). Most of the LUE-based models work at monthly or annual time scale, and rely on a linear relationship between GPP and Absorbed Photosynthetically Active Radiation (APAR) and on a parabolic effect of temperature. Conversely, the Prelued model replicates GPP at a daily time scale, based on a nonlinear relationship between APAR and GPP (Medlyn et al., 2003; Turner et al., 2003), a saturating effect of average daily temperature (which simulates the ecosystem acclimation to temperature, Mäkelä et al. (2004)), and daily meteorological and environmental variables. Such response to these environmental variables improves the fit of the model especially in temperature-controlled ecosystems (McCallum et al., 2013).

One of the critical aspects in the application of Prelued is the estimation of the model parameters and of the uncertainty associated with them. For this purpose we considered the application of the Bayesian model calibration, a method that has become more and more popular in

the last few years to obtain insights on both model predictions and uncertainties. The main characteristic of a Bayesian calibration is that it quantifies model inputs and outputs in the form of probability distributions, and applies the rules of probability theory to update the distributions when new data are obtained (Sivia, 1996; van Oijen et al., 2005). This approach has been widely used in different fields, and recently also to a large number of forest models with different structure and aims(Chevallier et al., 2006; Jansson et al., 2008; van Oijen et al., 2005, 2011). Even so, the application of the Bayesian method to LUE-based models is not as common as its application to process-based models, with a very few studies heading in this direction (Still et al., 2004; Xenakis et al., 2008).

In this study, we applied the Prelued model to eight italian Eddy-Covariance forest sites, with a Bayesian approach to calibration, and studied in detail the trend in the responses to environmental variables in each site to detect their importance in driving daily GPP. To our knowledge, this model has never been applied before to ecosystems characterized by such a wide range of climatic and environmental conditions. Moreover, since Prelued has never been calibrated following a Bayesian approach, before this study there was no information in the existing literature about uncertainties around the parameter values, nor about uncertainties on the model estimates. The aims of this work were therefore: 1) testing the ability of the Prelued model to simulate GPP at contrasting forest sites characterized by very different climates, elevations and plant functional types, 2) testing if the Bayesian approach improves the model performances in respect to other methods, and 3) estimating uncertainties around both parameter values and model estimates of daily GPP.

2. Materials and methods

2.1 Model formulation

The model used in this study was a modified version of Prelued, a LUE-type model of daily photosynthetic production of the canopy, developed by Mäkelä et al.(2008). Daily GPP is calculated as follows:

$$GPP_j = \beta APAR_j \prod_i F_{ij} \qquad (1)$$

where GPP$_j$ is canopy Gross Primary Production (gC m^{-2}) during day j, β is potential daily Light Use Efficiency (gC mol^{-1}), APAR$_j$ is Absorbed Photosynthetic Active Radiation (mol m^{-2}) during day j, and F$_{ij}$∈ [0, 1] are modifying factors accounting for suboptimal conditions in day j. The actual LUE of the canopy in day j is the product of β and the current values of the modifiers.

In the original version of the model, four modifiers were considered: a light modifier (FL) was defined so as to yield a rectangular hyperbola when multiplied with the linear response included in the LUE model, a temperature modifier (FS) was defined using the concept of state of acclimation (Mäkelä et al., 2004), a Vapour Pressure Deficit (VPD) modifier (FD) was defined following Landsberg and Waring (1997), and a Soil Water Content (SWC) modifier (FW) was based on the relative extractable water, dependent on wilting point and field capacity.

Since the wilting point is a difficult variable to estimate, FW was reformulated as follows to avoid the need of wilting point as an input variable:

$$FW_j = \begin{cases} 1 & SWC_j > \theta_{FC} \\ \left[1 + \left(\frac{\theta_{FC} - SWC_j}{\alpha'} \right)^\nu \right]^{-1} & SWC_j < \theta_{FC} \end{cases} \qquad (2)$$

where SWC$_j$ is volumetric Soil Water Content (SWC) (m^{3}m^{-3}), θ_{FC} is SWC at field capacity and α' is the new parameter used instead of α, thus $\alpha' = \alpha$ ($\theta_{FC}-\theta_{WP}$), where θ_{WP} is SWC at permanent wilting point (Mäkelä et al., 2008).

2.2 Data

This study used data collected with the eddy covariance (EC) technique at seven Italian forests sites and one Mediterranean macchia (table 1), downloaded from the European Fluxes Database Cluster (www.europe-fluxdata.eu). EC technique is a micrometeorological method that computes the net CO_2turbulent flux between a given ecosystem and the atmosphere from the covariance between the fluctuations of vertical wind velocity and CO_2 concentrations, averaged at a half-hour time scale. Currently, a global network of more than 500 EC stations exist worldwide to continuously monitor the CO_2 and energy exchange between ecosystems and the atmosphere,

whose homogeneity is ensured by similar standardizes procedures (Baldocchi, 2008). Collected data undergo quality tests and standardization procedures, following the approach proposed by Reichstein et al.(2005).

On the basis of data availability two one-year datasets were selected, one for the model calibration and one for the model validation. Daily average air temperature, relative humidity (Rh),SWC and daily total PAR were calculated from the database level 3 (thus quality assessed) hourly data using a very conservative approach (one missing data in the hourly time series for the day resulted in a missing daily value and therefore in a missing model output for that day)and used as input data for the model. Daily VPD was calculated from Rh and air temperature following Allen et al.(1998), and daily APAR was calculated using Normalized Difference Vegetation Index (NDVI) data as a proxy for the fraction of absorbed PAR (fAPAR),following Sims et al.(2006), as:

$$APAR_j = PAR_j (1.24 \, NDVI_j + 0.168) \qquad (3)$$

For this purpose, NDVI data with 0.25 km spatial grid and 16 days time-step were used (MODIS product MOD13Q1, 16-Day L3, 250m, Terra Vegetation Indices). To exclude days with NDVI equal to zero, which would result in null GPP no matter the other variables, a minimum threshold of 0.1 was set for NDVI (days with NDVI lower than 0.1 were considered missing) and no filter for cloudiness was applied.

Level 4 (thus quality assessed and gap-filled) daily GPP fluxes from the European Fluxes Database Cluster were used to calibrate the model. Missing data for either a weather variable or GPP resulted in a missing outcome of the model for that day j.

Field capacity for each site was estimated from soil texture data (Saxton et al., 1986).

The Bayesian calibration requires an estimate of the uncertainties around the data used during the calibration (van Oijen et al., 2005). Little is known in the literature about uncertainties on daily GPP, but Mo et al.(2008) estimated the uncertainties on average daily GPP to be15% and assumed them to be independent of each other. Uncertainties around daily GPP_j were therefore calculated as follows:

$$GPP_j = GPP_j \pm y_j \qquad (4)$$

where y_j randomly varied between 0.5 gC m^{-2} and 15% of daily GPP. The threshold for the minimum value was necessary to ensure that low values of GPP_j would not get an overwhelming weight during the calibration procedure.

2.3 Bayesian calibration

The prior parameter distributions were set based on the information made available by Mäkelä et al. (2008) and Peltoniemi et al.(2012). Since the parameter distributions were partly unknown, and since many parameters are empirical and without physiological meaning, we set the prior distributions as uniform (i.e. any value has the same probability to occur) and wide enough to cover a very wide range of possible values (table 2).

For each site, a 10^6 iterations long simulation was carried out using the Differential Evolution Markov Chain (DEMC) algorithm (ter Braak, 2006). An initial burn-in phase was set to 30% of the total number of iterations. 100 chains were considered, making the number of iterations per chain equal to 10^4.

A sample of parameter vectors from the posterior distribution generated by the Markov Chain Monte Carlo (MCMC) simulation was used to calculate the model results and uncertainties, and only the chain with maximum log-likelihood was sampled for this purpose. The number of candidate parameter sets in the sample is equal to 10% of the length of the Markov Chain, not considering the burn-in phase. The parameter set with maximum likelihood was used to calculate the model results, while the model uncertainties were calculated as the 5% and 95% percentile of the GPP distribution resulting from the model calculations for each parameter set in the MCMC sample.

2.4 Evaluation of model performance

Several indices were considered to evaluate model performances at each site. A linear regression of modeled vs. derived GPP values was fitted to calculate the coefficient of determination (R^2) and the root mean square error (RMSE). In addition, three other tests were

applied to evaluate the model performances. The first was Theil's inequality coefficient (U, Theil (1966)): Theil's U can assume values of 0 and greater. If $U = 0$ then the model produces perfect predictions. If $U = 1$ the model produces predictions of system behavior that are not better than a zero-change prediction. If $U > 1$, then the predictive power of the model is worse than the no-change prediction. The second test was the modelling efficiency (ME, Vanclay and Skovsgaard (1997)), which provides a simple index of performance on a relative scale: ME = 1 indicates a perfect fit, ME = 0 reveals that the model is no better than a simple average, while negative values indicate poor performance. The third test applied was the robust TOST test of equivalence with null hypothesis of dissimilarity of means (Robinson and Froese, 2004; Wellek, 2003): if the null hypothesis is rejected, the modeled and measured values belong to the same population, which indicates good model performance; vice versa, if the null hypothesis is not rejected, the modeled and measured values belong to different populations, which indicates poor model performance. Two criteria (based on two different ε values) were expressed relative to the sample standard deviation (25% and 50%) to represent a 'strict' and a 'liberal' criterion respectively, according to guidelines in Wellek (2003) and Blanco et al. (2007).

3. Results

3.1 Model performances

Model performances at the annual basis after calibration differed largely across sites. R^2 varied from 0.37 at the Mediterranean macchia site of Tolfa to 0.96 at the deciduous broadleaf forest of Collelongo, RMSE from 2.92 in Tolfa to 1.00at the sub-humid Mediterranean evergreen needleleaf forest of San Rossore, U from 0.14 in Collelongo and San Rossore to 0.38 in Tolfa, ME from 0.96 in Collelongo to 0.32 in Tolfa (table 3). The derived and modeled GPP were found to belong to the same population only when ε=0.5 in most sites, excluding Lavarone and Roccarespampani 2 (where the two datasets belong to different populations at all values of ε) and San Rossore (where the two datasets belong to the same populations at all values of ε). Overall, the model performed well in coniferous forests and in sites with climate typical of mountain regions, poorly in the Mediterranean

macchia, while its performance varied widely in broadleaf forests. The model always underestimated the average daily GPP (table 4, figure 1), with the strongest underestimation at Tolfa (1.35gC m^{-2} d^{-1}loweron average than the derived GPP), especially in the summer period (figure 2). In general, the model overestimated GPP at lower values of derived GPP, while underestimated GPP at higher values of derived GPP in all sites except at the mountain forests of Collelongo and Renon, where the model underestimated GPP even at low values of derived GPP (figure 3).

Model performances on the annual basis varied largely across sites also in validation. At Tolfa the model is unable to reproduce the GPP pattern (R^2=0.02, RMSE 2.78), confirming the model limitations in the simulation of this Mediterranean macchia. Concerning the other sites, R^2 varied from 0.49 in Castelporziano to 0.84 in Lavarone(both evergreen forests) and Nonantola, with RMSE ranging from 2.64 in Roccarespampani 2(the latter two being deciduous broadleaf forests) to 1.15 in San Rossore, U from 0.17 in San Rossore to 0.51 in Castelporziano, ME from 0.84 in Lavarone to -0.18 in Tolfa (table 3). The derived and modeled GPP were found to belong to the same population only when ε=0.5 and only in Lavarone and San Rossore. Overall, the model overestimated average daily GPP in Collelongo, Lavarone and San Rossore, while underestimated GPP in the other sites (table4, figure 1).The strongest overestimation occurred in Collelongo (1.47gC m^{-2} d^{-1} more than the average measured data), mainly during the autumn while the strongest underestimation occurred in Renon (1.43gC m^{-2} d^{-1} less than the data), where the model underestimated daily GPP all year (figure 4). As during calibration, the model overestimated GPP at lower values of derived GPP, while underestimated GPP at higher values of derived GPP in all sites except Renon (figure 5).

3.2 Parameter values and modifiers

The values and uncertainties of parameters are listed in table 5. The values of all the parameters were largely variable across sites, with no pattern related to climate, elevation or plant functional types (PFT), with the exception on the potential LUE (β) which ranged from 0.364gC mol^{-1} in

Collelongo to 1.495gC mol^{-1} in Nonantola (both deciduous broadleaf forests), generally with lower values in evergreen forests and higher values in the Mediterranean and in deciduous sites, with the exception of Collelongo.

Comparing the values of each modifier during the period of simulation was useful in order to understand the impact of each environmental factor on the modeled GPP (figure 6). In this study, the SWC modifier was always equal to 1 in Lavarone, Renon and Roccarespampani 2, indicating the absence of GPP limitation by drought stress in the modeled years, while was always below 0.92 in the Mediterranean deciduous broadleaf forest of Nonantola, indicating a continuous limitation of GPP due to water availability in the considered year. In the Mediterranean deciduous broadleaf forest of Roccarespampani 2, also light and VPD modifiers were always equal to 1, while the temperature modifier was always below 0.33, indicating a strong limitation of temperature all year on GPP and no limitation due to other factors. Continuous limitation due to temperature also occurred in the Mediterranean broadleaf forest in Castelporziano (temperature modifier always < 0.80) even if not as strong as in Roccarespampani 2, and in the Mediterranean deciduous broadleaf forest in Nonantola, where only in a few summer days the temperature modifier equaled 1.

3.3 Model uncertainties

For all the sites considered, both in calibration and validation, model uncertainties around daily GPP were lower than data uncertainties (figure 1), with the great majority of values between 0 and 1 gC m^{-2} d^{-1} in the calibration process, and extreme cases in Renon and Collelongo where uncertainties around the modeled GPP have values equal to 0 in several days. The difference between data and model uncertainties is higher in sites characterized by Mediterranean climate, with the exception of Renon during validation.

4. Discussion

4.1 Model performance and evaluation

The model had similar good performances as reported by Mäkelä et al.(2008) and Peltoniemi et al. (2012) in coniferous forests and in sites with climate typical of mountain regions, while the

performances varied considerably across broadleaf forests in a way that appeared to be independent of the type of climate, and were not satisfactory in the Mediterranean macchia. The last result was expected since LUE models are likely to have difficulties in estimating GPP in ecosystems where productivity is limited by water availability: this has been pointed out by Schaefer et al. (2012) and could be due to equifinality between VPD and SWC modifiers, since the effects of drought and humidity stress are strictly associated. In the case of Prelued, the SWC modifier was found to be not significant by Mäkelä et al.(2008) for the coniferous sites of Sodankylä (Finland), Hyytiälä (Finland), Norunda (Sweden), Tharandt (Germany) and Bray (France), while it seemed to be an important factor in many sites we considered (figure 6). A large variability in model performances across regions and sites was also found by Peltoniemi et al.(2012) and McCallum et al. (2013), and was probably due to spatial changes in the maximum LUE parameter. In our case, this variability is likely due to the very heterogeneous conditions and characteristics of the sites considered: in particular, the model had difficulties in reproducing GPP in drought-limited sites (as in Schaefer et al. (2012)), characterized by a Mediterranean climate, conditions very different from the ones that can be found in boreal, less drought-limited sites where the model has proven to perform best (Mäkelä et al., 2008).

An interesting exception to the good performance of the model in temperate coniferous sites was Renon, where the model performances were poor during validation, with a strong underestimation during the entire considered year. Poor model performance in this particular site was also observed by Peltoniemi et al. (2012). These authors considered the high elevation and the unaccounted CO_2 advected flux (Marcolla et al., 2005; Montagnani et al., 2009)as the main causes, but many other factors could also be considered. Renon is an open forest with many gaps and the presence of grass around and inside the tower footprint alters the NDVI retrievals from the satellite (NDVI ranges from 0.16 to 0.76 in 2005 and from 0.10 to 0.84 in 2010), finally affecting the estimate of APAR and causing the model to underestimate GPP. Uncertainties in the estimate of APAR related to ground vegetation and of the fAPAR has been pointed out as one of the most

important factors affecting the model results also by Mäkelä et al.(2008). Another possible explanation, related to the differences in the model performances between calibration and validation, was that the application of Prelued in two different years relies on the assumption that the sites characteristics (e.g. potential LUE, field capacity) remain constant over time, which could be a weak assumption in this particular case.

The model performed generally well also in deciduous forests (e.g. Collelongo and Nonantola), but seemed unable to reproduce the peak in GPP occurring in the spring (figure 2). There are several possible explanations for this behavior, the most important related to the model response to temperature and phenology. The acclimation function for the temperature modifier was developed and tested on evergreen coniferous forests, but may not be appropriate to describe the physiological processes associated with the restart of the physiological activities in deciduous forests. The phenological cycle is taken into account only via the calculation of APAR through fAPAR or NDVI, but it is well known to depend on both temperature and radiation (Häkkinen et al., 1998; Melaas et al., 2013). Therefore, NDVI or fAPAR alone are not sufficient to describe phenology, especially at sites with dense understorey vegetation (e.g. Nonantola) or with large forest gaps (e.g. Renon).

Potential LUE values (parameter β) in this study were highly variable between sites as in Mäkelä et al. (2008), and in accordance with Turner et al. (2003) and Kergoat et al. (2008), who pointed out that the maximum LUE seems to vary by PFT. Compared to Mäkelä et al. (2008), our estimates in coniferous forests were at their lower range for the same kind of ecosystems, while comparable with those found Peltoniemi et al. (2012). As in the latter study, we found lower potential LUE values generally associated with evergreen forests and higher potential LUE associated with deciduous forests. This result is consistent with several other studies (e.g. Gower et al., 1999; Turner et al., 2003; Yuan et al., 2007), follows the known ecological pattern of LUE in different biomes and climate conditions, and can be explained by ample thermal resources and moisture in deciduous forests. No association was found between the remaining parameters related

to the response to meteorological drivers and climate, elevation nor PFT, which was surprising considering the wide range of conditions across sites.

Even if Prelued is a simplified empirical model, the annual trends of the modifiers gave the opportunity to evaluate the model behavior from a more ecological point of view, allowing us to investigate the temporal dynamic of the relative importance of each climatic driver, and providing a tool to detect if and where the model performances were good for the wrong reasons. In this study, several interesting points emerged from this analysis (figure6). As expected, the soil water content acted as a limiting factor in sites characterized by a Mediterranean climate, particularly at Castelporziano, Nonantola, San Rossore and Tolfa, with a stronger impact on daily GPP starting from spring. This was in accordance with Yuan et al.(2007), who found that in Mediterranean climate, soil moisture was the dominant control factor during the growing season. On the contrary, the soil water content was never a limiting factor at Renon and Lavarone, thanks to the sub-humid temperate mountain climate of these sites, which did not experience periods of drought in the considered years.

If the response to soil water content was in accordance with the scientific literature and the known ecological patterns, on the other hand the analysis of the temperature modifier showed a less reliable behavior at the Mediterranean broadleaf forest in Roccarespampani 2. At this site, the model predicted a marked temperature limitation throughout the year, while the other variables were considered always non-limiting. A similar pattern was observed for the Mediterranean sites of Castelporziano and Nonantola, even if the limitation in GPP due to temperature was not as extreme as in Roccarespampani 2. Given the characteristics of these sites (low elevation and Mediterranean climate), this result has no ecological explanation, and represents a major problem in evaluating the model behavior and performance. To further investigate the problem, three calibrations were run for this site with the same results (data not shown), which are likely due to the MCMC simulation to fail reaching a proper convergence despite the use of a complex algorithm with a very high number of iterations. This failure was possibly caused by the mathematical structure of Prelued: the

multiplicative structure could generate very high correlations between parameters, allowing a single modifier to account for a great part of the data variability without impairing the model capability to fit to the data (Mäkelä et al., 2008). These correlations reflect the interdependence between environmental variables (i.e. PAR and the driving variables in the modifiers, or mean VPD and LUE), as found by Peltoniemi et al. (2012). The temperature modifier raised some concerns also in Mäkelä et al. (2008), where the temperature was considered non-limiting most of the time in Bray, and the model results were insensitive to any of the related parameters. It is also well established that VPD and temperature are tightly linked, which could cause problems related to equifinality.

4.2 Model and parameter uncertainties

The Bayesian approach to calibration allowed us to study both parameter and model uncertainties in detail even if the variability among sites and the ability of Prelued to fit the data did not confer a clear advantage to the Bayesian calibration in terms of model performances. The model uncertainties were very small when compared to the uncertainties of the derived GPP, possibly because being a simplified empirical model, Prelued does not consider all the possible sources of variability in the derived GPP, likely missing one or more of them, which cause the model uncertainties to be lower than data uncertainties. Peltoniemi et al. (2012) pointed out that some variation not captured by the model may be due to systematic features in the EC measurements not taken into account by the model. Some examples are the variation of footprint area due to wind direction and turbulence regime (Rannik et al., 2006), the variation in the contribution from trees, ground vegetation and soil to CO_2 fluxes (Davidson et al., 2006), and the error in the estimate of total ecosystem respiration (Mäkelä et al., 2006). Another possible explanation was related to the calculation of uncertainties around the derived GPP, which were taken into account in the calibration process to estimate the log-likelihood of each candidate parameter vector. Very few examples can be found in the literature related to estimates of uncertainties around daily GPP: based on Mo et al.(2008) we estimated the maximum uncertainties around daily derived GPP to be 15% of its value, while Duursma et al. (2009)estimated the daily uncertainties to be around 5% of

the daily GPP. There is no well-established procedure to estimate these uncertainties, but given their importance in the modeling process their value is paramount for a correct estimate of uncertainties on model results (Raupach et al., 2005; Richardson and Hollinger, 2005). Both the lack of a well-established procedure and the complexity of the environmental background of daily GPP make the probabilistic Bayesian approach used in this study an ideal solution to estimate the uncertainties on modeled daily GPP. Most published sensitivity analyses rely on a local approach, giving information of the influence of one input factor variability on model output, keeping all other factors constant ('one-at-a-time' approach): this kind of approach has been proven generally inadequate by Saltelli and Annoni(2010), given that it 1) can be highly biased for non linear systems in which the input factors are highly interactive, and 2) relies on the assumptions of model linearity. On the other hand, a sensitivity analysis with global approach (like Global Sensitivity and Uncertainty Analysis) would take correlations between input factors into account, but is usually applied to very complex models (Convertino et al., 2014) and would be inadequate considering the simplicity of Prelued, which does not include (on purpose) several variables and processes known to play a role in determining the global productivity of an ecosystem.

The uncertainties around the parameter values were cause of some concerns: the uncertainties on potential LUE (β)for the Prelued model were estimated by Peltoniemi et al. (2012) to be in the range of 0.1-0.2 gC m^{-2} d^{-1}(standard error) both for coniferous and deciduous forests. In our results, the only sites in the same range were the alpine coniferous sites of Lavarone and Renon, which experience a similar climate as boreal coniferous forests where the model had been tested before our work, and in the Mediterranean site of Nonantola. Uncertainties on β were around half the expected range in San Rossore (Mediterranean coniferous forest) and in Collelongo (mountain Mediterranean deciduous forest), and much larger in Castelporziano and Roccarespampani 2, probably due to difficulties in reaching convergence as mentioned in section 4.2.

No comparison could be done with the existing literature concerning uncertainties on the parameters related to response to temperature, due to the calibration method used in Mäkelä et

al.(2008) and in Peltoniemi et al. (2012), the latter showing only uncertainties around potential LUE. No parameter uncertainties were shown in McCallum et al (2013).

The uncertainties were remarkably variable for parameters other than β, and very large on parameters related to the response to SWC in many sites. Together with the difficulties in estimating reliable parameter values in Roccarespampani 2 and Castelporziano, this raised some concerns about the reaching of proper convergence by the MCMC during the calibration process, which could be due to the multiplicative structure of Prelued(Mäkelä et al., 2008).

5. Conclusions

We applied the Prelued model to eight Italian eddy-covariance sites, over contrasting natural vegetation types that represented a wide range of climatic and environmental conditions. The model performed well in almost all sites considered with the exception of a Mediterranean macchia: therefore, it could be successfully applied to different forest ecosystems, similar or different from the ones considered in this study. However, the issues related to its mathematical structure and the model performances during validation in several sites suggest a careful site-specific calibration on the period of simulation, and the model responses to driving climatic variables should be carefully taken into account to evaluate its ecological behavior (e.g. model fitting in three Mediterranean sites was the result of unrealistic responses to climatic variables).

Even if the Bayesian approach to calibration did not bring an evident advantage concerning model performances, it allowed us to estimate uncertainties around all parameter values, highlighting some possible convergence issue during the calibration process, and uncertainties around model estimates of daily GPP, which resulted to be always smaller than uncertainties around the data. An improvement of the estimates of uncertainties around daily GPP derived from eddy-covariance measurements, together with more information on the prior probability distributions of the model parameters and the application of different Monte Carlo Markov Chain algorithms, would help to clarify these issues.

6. References

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Trans. Automat. Contr. 19, 716–723.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements. Irrig. Drain. Pap. 65 (300 pp.), United Nations Food Agric. Organ. Rome, Italy.

Andrieu, C., de Freitas, N., Doucet, A., Jordan, M., 2003. An Introduction to MCMC for Machine Learning. Mach. Learn. 50, 5–43. doi:10.1023/A:1020281327116

Asmussen, S., Glynn, P.W., 2007. Stochastic simulation: Algorithms and analysis. Springer Science & Business Media.

Baldocchi, D.D., 2008. "Breathing" of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. 56, 1–26. doi:10.1071/BT07151

Bernardo, J.M., Smith, A.F.M., 2009. Bayesian theory. John Wiley & Sons.

Blanco, J.A., Seely, B., Welham, C., Kimmins, J.P. (Hamish), Seebacher, T.M., 2007. Testing the performance of a forest ecosystem model (FORECAST) against 29 years of field data in a Pseudotsuga menziesii plantation. Can. J. For. Res. 37, 1808–1820. doi:10.1139/X07-041

Bloom, A.A., Williams, M., 2015. Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological" common sense" in a model–data fusion framework. Biogeosciences 12, 1299–1315.

Browne, W.J., Steele, F., Golalizadeh, M., Green, M.J., 2009. The use of simple reparameterizations to improve the efficiency of Markov chain Monte Carlo estimation for multilevel models with applications to discrete time survival models. J. R. Stat. Soc. Ser. A (Statistics Soc. 172, 579–598.

Bryson, A.E., 1975. Applied optimal control: optimization, estimation and control. CRC Press.

Buzzi-Ferraris, G., Manenti, F., 2010. Better reformulation of kinetic models. Comput. Chem. Eng. 34, 1904–1906.

Canadell, J.G., Le Quéré, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A., Marland, G., 2007. Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. Proc. Natl. Acad. Sci. 104 , 18866–18870. doi:10.1073/pnas.0702737104

Casella, G., Robert, C.P., 1999. Monte Carlo statistical methods.

Chevallier, F., Viovy, N., Reichstein, M., Ciais, P., 2006a. On the assignment of prior errors in Bayesian inversions of CO2 surface fluxes. Geophys. Res. Lett. 33, 1–5. doi:10.1029/2006GL026496

Chevallier, F., Viovy, N., Reichstein, M., Ciais, P., 2006b. On the assignment of prior errors in Bayesian inversions of CO2 surface fluxes. Geophys. Res. Lett. 33, 1–5. doi:10.1029/2006GL026496

Convertino, M., Muñoz-Carpena, R., Chu-Agor, M.L., Kiker, G.A., Linkov, I., 2014. Untangling drivers of species distributions: Global sensitivity and uncertainty analyses of MAXENT. Environ. Model. Softw. 51, 296–309. doi:10.1016/j.envsoft.2013.10.001

Davi, H., Dufrene, E., Francois, C., Lemaire, G., Loustau, D., Bosc, a, Rambal, S., Granier, a, Moors, E., 2006. Sensitivity of water and carbon fluxes to climate changes from 1960 to 2100 in European forest ecosystems. Agric. For. Meteorol. 141, 35–56.

Davidson, E.A., Richardson, A.D., Savage, K.E., Hollinger, D.Y., 2006. A distinct seasonal pattern of the ratio of soil respiration to total ecosystem respiration in a spruce-dominated forest. Glob. Chang. Biol. 12, 230–239. doi:0.1111/j.1365-2486.2005.01062.x

De Weirdt, M., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., Ciais, P., Steppe, K., 2012. Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model. Geosci. Model Dev. 5, 1091–1108. doi:10.5194/gmd-5-1091-2012

Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C., Wisniewski, J., 1994. Carbon pools and flux of global forest ecosystems. Science (80-.). 263, 185–190.

Dufrêne, E., Davi, H., François, C., Maire, G. Le, Dantec, V. Le, Granier, a., 2005. Modelling carbon and water cycles in a beech forestPart I: Model description and uncertainty analysis on modelled NEE. Ecol. Modell. 185, 407–436.

Duursma, R. a, Kolari, P., Perämäki, M., Pulkkinen, M., Mäkelä, A., Nikinmaa, E., Hari, P., Aurela, M., Berbigier, P., Bernhofer, C.H., Grünwald, T., Loustau, D., Mölder, 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 Physiol. 29, 621–639. doi:10.1093/treephys/tpp010

Ellison, A.M., 2004. Bayesian inference in ecology. Ecol. Lett. 7, 509–520.

Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecol. Appl. 6, 1036–1046.

Field, C.B., 2001. Plant Physiology of the "Missing" Carbon Sink. Plant Physiol. 125, 25–28.

Foley, J.A., Prentice, I.C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., Haxeltine, A., 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochem. Cycles 10, 603–628. doi:10.1029/96GB02692

Franks, S.W., Beven, K.J., 1997. Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. J. Geophys. Res. 102, 23,991–23,999.

Fu, Y.H., Campioli, M., Van Oijen, M., Deckmyn, G., Janssens, I.A., 2012. Bayesian comparison of six different temperature-based budburst models for four temperate tree species. Ecol. Modell. 230, 92–100.

Gao, C., Wang, H., Weng, E., Lakshmivarahan, S., Zhang, Y., Luo, Y., 2011. Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics. Ecol. Appl. 21, 1461–1473.

Ghazoul, J., McAllister, M., 2003. Communicating complexity and uncertainty in decision-making contexts: Bayesian approaches to forest research. Int. For. Rev. 5, 9–19.

Gilks, W., Roberts, G., 1996. Strategies for Improving MCMC., in: Gilks, W., Richardson, S., Spiegelhalter, D. (Eds.), Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, FL., pp. 89–114.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov chain Monte Carlo in practice. CRC press.

Goodale, C.L., Apps, M.J., Birdsey, R.A., Field, C.B., Heath, L.S., Houghton, R.A., Jenkins, J.C., Kohlmaier, G.H., Kurz, W., Liu, S., Nabuurs, G.-J., Nilsson, S., Shvidenko, A.Z., 2002. Forest Carbon Sinks In The Northern Hemisphere. Ecol. Appl. 12, 891–899. doi:10.1890/1051-0761(2002)012[0891:FCSITN]2.0.CO;2

Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote Sens. Environ. 70, 29–51. doi:10.1016/S0034-4257(99)00056-5

Grant, R.F., Baldocchi, D.D., Ma, S., 2012. Ecological controls on net ecosystem productivity of a seasonally dry annual grassland under current and future climates: Modelling with ecosys. Agric. For. Meteorol. 152, 189–200.

Grassi, G., Magnani, F., 2005. Stomatal, mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees. Plant, Cell Environ. 28, 834–849. doi:10.1111/j.1365-3040.2005.01333.x

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli 7, 223–242.

Häkkinen, R., Linkosalo, T., Hari, P., 1998. Effects of dormancy and environmental factors on timing of bud burst in Betula pendula. Tree Physiol. 18, 707–712. doi:10.1093/treephys/18.10.707

Hill, T.C., Ryan, E., Williams, M., 2012. The use of CO2 flux time series for parameter and carbon stock estimation in carbon cycle research. Glob. Chang. Biol. 18, 179–193.

Hollinger, D.Y., Richardson, A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol. 25, 873–885. doi:10.1093/treephys/25.7.873

Horn, J.E., Schulz, K., 2011. Identification of a general light use efficiency model for gross primary production. Biogeosciences 8, 999–1021. doi:10.5194/bg-8-999-2011

Horn, J.E., Schulz, K., 2011. Spatial extrapolation of light use efficiency model parameters to predict gross primary production. J. Adv. Model. Earth Syst. 3, 21:

Huntington, T.G., 2006. Evidence for intensification of the global water cycle: Review and synthesis. J. Hydrol. 319, 83–95. doi:http://dx.doi.org/10.1016/j.jhydrol.2005.07.003

Jansson, P.-E., Svensson, M., Kleja, D.B., Gustafsson, D., 2008. Simulated climate change impacts on fluxes of carbon in Norway spruce ecosystems along a climatic transect in Sweden. Biogeochemistry 89, 81–94.

Jarvis, A.J., Stauch, V.J., Schulz, K., Young, P.C., 2004. The seasonal temperature dependency of photosynthesis and respiration in two deciduous forests. Glob. Chang. Biol. 10, 939–950. doi:10.1111/j.1529-8817.2003.00743.x.

Jaynes, E.T., 2003. Probability theory: the logic of science. Cambridge university press.

Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., Saugier, B., 2008. Nitrogen controls plant canopy light-use efficiency in temperate and boreal ecosystems. J. Geophys. Res. 113, 1–19. doi:10.1029/2007JG000676

Kimmins, J.P., 1986. Predicting the consequences of intensive forest harvesting on long-term productivity: the need for a hybrid model such as FORCYTE-11. Rapp. Institutionen Ekol. och Miljovard, Sveriges Lantbruksuniversitet 31–84.

Kimmins, J.P., Blanco, J.A., Seely, B., Welham, C., Scoullar, K., 2008. Complexity in modelling forest ecosystems: How much is enough? For. Ecol. Manage. 256, 1646–1658. doi:10.1016/j.foreco.2008.03.011

Kimmins, J.P., Mailly, D., Seely, B., 1999. Modelling forest ecosystem net primary production: the hybrid simulation approach used in forecast. Ecol. Modell. 122, 195–224. doi:10.1016/S0304-3800(99)00138-6

Kramer, K., Leinonen, I., Bartelink, H.H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, a. J., Froer, O., Gracia, C. a., Granier, a., Grunwald, T., Hari, P., Jans, W., Kellomaki, S., Loustau, D., Magnani, F., Markkanen, T., Matteucci, G., Mohren, G.M.J., Moors, E., Nissinen, a., Peltola, H., Sabate, S., Sanchez, a., Sontag, M., Valentini, R., Vesala, T., 2002. Evaluation of six process-based forest growth models using eddy-covariance measurements of CO2 and H2O fluxes at six forest sites in Europe. Glob. Chang. Biol. 8, 213–230. doi:10.1046/j.1365-2486.2002.00471.x

Kumar, M., Monteith, J.L., 1981. Remote sensing of crop growth, in: Plants and the Daylight Spectrum. Academic Press: London, pp. 133–144.

Landsberg, J., 2003. Modelling forest ecosystems: state of the art, challenges, and future directions. Can. J. For. Res. 33, 385–397. doi:10.1139/x02-129

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

Landsberg, J., Waring, R.H., Coops, N.C., 2003. Performance of the forest productivity model 3-PG applied to a wide range of forest types. For. Ecol. Manage. 172, 199–214. doi:10.1016/S0378-1127(01)00804-0

Magnani, F., Grace, J., Borghetti, M., 2002. Adjustment of tree structure in response to the environment under hydraulic constraints. Funct. Ecol. 16, 385–393. doi:10.1046/j.1365-2435.2002.00630.x

Mäkelä, A., 1988. Performance analysis of a process-based stand growth model using Monte Carlo techniques. Scand. J. For. Res. 3, 315–331.

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, 369–76. doi:10.1093/treephys/24.4.369

Mäkelä, A., Kolari, P., Karimäki, J., Nikinmaa, E., Perämäki, M., Hari, P., 2006. Modelling five years of weather-driven variation of GPP in a boreal forest. Agric. For. Meteorol. 139, 382–398. doi:0.1016/j.agrformet.2006.08.017

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen, P., 2000a. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20, 289–298.

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen, P., 2000b. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20, 289–298.

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. Chang. Biol. 14, 92–108. doi:10.1111/j.1365-2486.2007.01463.x

Marcolla, B., Cescatti, A., Montagnani, L., Manca, G., Kerschbaumer, G., Minerbi, S., 2005. Importance of advection in the atmospheric CO2 exchanges of an alpine forest. Agric. For. Meteorol. 130, 193–206. doi:10.1016/j.agrformet.2005.03.006

McCallum, I., Franklin, O., Moltchanova, E., Merbold, L., Schmullius, C., Shvidenko, A., Schepaschenko, D., Fritz, S., 2013. Improved light and temperature responses for light-use-efficiency-based GPP models. Biogeosciences 10, 6577–6590. doi:10.5194/bg-10-6577-2013

McMurtrie, R.E., Gholz, H.L., Linder, S., Gower, S.T., 1994. Climatic factors controlling the productivity of pine stands : a model-based analysis. Ecol. Bull. 43, 173–188.

Medlyn, B., Barrett, D., Landsberg, J., Sands, P., Clement, R., 2003. Conversion of canopy intercepted radiation to photosynthate: review of modelling approaches for regional scales. Funct. Plant Biol. 30, 153–169. doi:10.1071/FP02088

Melaas, E.K., Richardson, A.D., Friedl, M.A., Dragoni, D., Gough, C.M., Herbst, M., Montagnani, L., Moors, E., 2013. Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. Agric. For. Meteorol. 171-172, 46–56. doi:110.1016/j.agrformet.2012.11.018

Milne, A.E., Glendining, M.J., Bellamy, P., Misselbrook, T., Gilhespy, S., Casado, M.R., Hulin, A., Van Oijen, M., Whitmore, A.P., 2014. Analysis of uncertainties in the estimates of nitrous

oxide and methane emissions in the UK's greenhouse gas inventory for agriculture. Atmos. Environ. 82, 94–105.

Minunno, F., Xenakis, G., Perks, M.P., Mencuccini, M., 2010. Calibration and validation of a simplified process- based model for the prediction of the carbon balance of Scottish Sitka spruce (Picea sitchensis) plantations. Can. J. For. Res. 40, 2411–2426. doi:10.1139/X10-181

Mo, X., Chen, J.M., Ju, W., Black, T.A., 2008. Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an ensemble Kalman filter. Ecol. Modell. 217, 157–173. doi:10.1016/j.ecolmodel.2008.06.021

Montagnani, L., Manca, G., Canepa, E., Georgieva, E., Acosta, M., Feigenwinter, C., Janous, D., Kerschbaumer, G., Lindroth, A., Minach, L., Minerbi, S., Mölder, M., Pavelka, M., Seufert, G., Zeri, M., Ziegler, W., 2009. A new mass conservation approach to the study of CO2 advection in an alpine forest. J. Geophys. Res. Atmos. 114, D07306. doi:10.1029/2008JD010650

Monteith, J.L., Moss, C.J., 1977. Climate and the Efficiency of Crop Production in Britain [and Discussion]. Philos. Trans. R. Soc. Lond. B 281, 277–294. doi:10.1098/rstb.1977.0140

Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., Liski, J., Zhou, L., Alexeyev, V., Hughes, M.K., 2001. A large carbon sink in the woody biomass of Northern forests. Proc. Natl. Acad. Sci. 98 , 14784–14789. doi:10.1073/pnas.261555198

Nightingale, J.M., Phinn, S.R., Held, A.A., 2004. Ecosystem process models at multiple scales for mapping tropical forest productivity. Prog. Phys. Geogr. 28, 241–281. doi:10.1191/0309133304pp411ra

Ogle, K., 2009. Hierarchical Bayesian statistics: merging experimental and modeling approaches in ecology. Ecol. Appl. 19, 577–581.

Ogle, K., Barber, J.J., 2008. Bayesian data—model integration in plant physiological and ecosystem ecology, in: Progress in Botany. Springer, pp. 281–311.

Oulehle, F., Evans, C.D., Hofmeister, J., Krejci, R., Tahovska, K., Persson, T., Cudlin, P., Hruska, J., 2011. Major changes in forest carbon and nitrogen cycling caused by declining sulphur deposition. Glob. Chang. Biol. n/a–n/a. doi:10.1111/j.1365-2486.2011.02468.x

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., 2012a. Does canopy mean nitrogen concentration explain variation in canopy light use efficiency across 14 contrasting forest sites? Tree Physiol. 32, 200–18. doi:10.1093/treephys/tpr140

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

Pretzsch, H., Grote, R., Reineking, B., Rötzer, T., Seifert, S., 2008. Models for forest ecosystem management: a European perspective. Ann. Bot. 101, 1065–1087. doi:10.1093/aob/mcm246

Rahn, K., Butterbach-bahl, K., Werner, C., 2011. Selection of likelihood parameters for complex models determines the effectiveness of Bayesian calibration. Ecol. Inform. 6, 333–340. doi:10.1016/j.ecoinf.2011.08.002

Rannala, B., 2002. Identifiability of Parameters in MCMC Bayesian Inference of Phylogeny. Syst. Biol. 51, 754–760. doi:10.1080/10635150290102429

Rannik, Ü., Kolari, P., Vesala, T., Hari, P., 2006. Uncertainties in measurement and modelling of net ecosystem exchange of a forest. Agric. For. Meteorol. 138, 244–257. doi:10.1016/j.agrformet.2006.05.007

Rastetter, E.B., Aber, J.D., Peters, D.P.C., Ojima, D.S., Burke, I.C., 2003. Using mechanistic models to scale ecological processes across space and time. Bioscience 53, 68–76.

Raupach, M.R., Rayner, P.J., Barrett, D.J., DeFries, R.S., Heimann, M., Ojima, D.S., Quegan, S., Schmullius, C.C., 2005. Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Glob. Chang. Biol. 11, 378–397. doi:10.1111/j.1365-2486.2005.00917.x

Reckhow, K.H., 1990. Bayesian inference in non-replicated ecological studies. Ecology 71, 2053–2059.

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. Chang. Biol. 11, 1424–1439. doi:10.1111/j.1365-2486.2005.001002.x

Richardson, A.D., Hollinger, D.Y., 2005. Statistical modeling of ecosystem respiration using eddy covariance data: Maximum likelihood parameter estimation, and Monte Carlo simulation of model and parameter uncertainty, applied to three simple models. Agric. For. Meteorol. 131, 191–208. doi:10.1016/j.agrformet.2005.05.008

Robinson, A.P., Froese, R.E., 2004. Model validation using equivalence tests. Ecol. Modell. 176, 349–358. doi:10.1016/j.ecolmodel.2004.01.013

Rodeghiero, M., Cescatti, A., 2005. Main determinants of forest soil respiration along an elevation/temperature gradient in the Italian Alps. Glob. Chang. Biol. 11, 1024–1041. doi:10.1111/j.1365-2486.2005.00963.x

Running, S.W., Baldocchi, D.D., Turner, D.P., Gower, S.T., Bakwin, P.S., Hibbard, K.A., 1999. A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sens. Environ. 70, 108–127.

Sacks, W.J., Schimel, D.S., Monson, R.K., Braswell, B.H., 2006. Model-data synthesis of diurnal and seasonal CO2 fluxes at Niwot Ridge, Colorado. Glob. Chang. Biol. 12, 240–259. doi:10.1111/j.1365-2486.2005.01059.x

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environ. Model. Softw. 25, 1508–1517. doi:10.1016/j.envsoft.2010.04.012

Saxton, K.E., Rawls, W.J., Romberger, J.S., Papendick, R.I., 1986. Estimating Generalized Soil-water Characteristics from Texture1. Soil Sci. Soc. Am. J. 50, 1031–1036. doi:10.2136/sssaj1986.03615995005000040039x

Schaefer, K., Schwalm, C.R., Williams, C., Arain, M.A., Barr, A., Chen, J.M., Davis, K.J., Dimitrov, D., Hilton, T.W., Hollinger, D.Y., 2012. A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. J. Geophys. Res. 117. doi:10.1029/2012JG001960

Schulz, K., Jarvis, A.J., Beven, K.J., Soegaard, H., 2001. The Predictive Uncertainty of Land Surface Fluxes in Response to Increasing Ambient Carbon Dioxide. J. Clim. 14, 2551–2562. doi:http://dx.doi.org/10.1175/1520-0442(2001)014<2551:TPUOLS>2.0.CO;2

Sims, D.A., Luo, H., Hastings, S., Oechel, W.C., Rahman, A.F., Gamon, J.A., 2006. Parallel adjustments in vegetation greenness and ecosystem CO2 exchange in response to drought in a Southern California chaparral ecosystem. Remote Sens. Environ. 103, 289–303. doi:10.1016/j.rse.2005.01.020

Sivia, D.S., 1996. Data Analysis: A Bayesian Tutorial. Oxford University Press, Oxford.

Smith, T.J., Marshall, L.A., 2008. Bayesian methods in hydrologic modeling : A study of recent advancements in Markov chain Monte Carlo techniques 44, 1–9. doi:10.1029/2007WR006705

Still, C.J., Randerson, J.T., Fung, I.Y., 2004. Large-scale plant light-use efficiency inferred from the seasonal cycle of atmospheric CO2. Glob. Chang. Biol. 10, 1240–1252. doi:10.1111/j.1365-2486.2004.00802.x

Storn, R., Price, K., 1997. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim. 11, 341–359. doi:10.1023/A:1008202821328

Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., Lindroth, a, 2008a. Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand. Ecol. Modell. 213, 331–344. doi:10.1016/j.ecolmodel.2008.01.001

Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., Lindroth, A., 2008b. Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand. Ecol. Modell. 213, 331–344. doi:10.1016/j.ecolmodel.2008.01.001

Ter Braak, C.J.F., 2006. A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. Stat. Comput. 16, 239–249. doi:10.1007/ s11222-006-8769-1

Theil, H., 1966. Applied Econometric Forecasting. North-Holland, Amsterdam.

Tjiputra, J., Roelandt, C., Bentsen, M., Lawrence, D., Lorentzen, T., Schwinger, J., Seland, O., Heinze, C., 2013. Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). Geosci. Model Dev. 6, 301–325. doi:10.5194/gmd-6-301-2013

Turner, D.P., Ritts, W.D., Styles, J.M., Yang, Z., Cohen, W.B., Law, B.E., Thornton, P.E., 2006. A diagnostic carbon flux model to monitor the effects of disturbance and interannual variation in climate on regional NEP. Tellus B 58, 476–490. doi:10.1111/j.1600-0889.2006.00221.x

Turner, D.P., Urbanski, S., Bremer, D., Wofsy, S.C., Meyers, T., Gower, S.T., Gregory, M., 2003. A cross-biome comparison of daily light use efficiency for gross primary production. Glob. Chang. Biol. 9, 383–395. doi:10.1046/j.1365-2486.2003.00573.x

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. doi:10.1016/j.agrformet.2011.06.017

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. doi:10.1016/j.agrformet.2011.06.017

Van Oijen, M., Cannell, M.G.R., Levy, P.E., 2004. Modelling biogeochemical cycles in forests: state of the art and perspectives, in: Andersson, F., Birot, Y., Päivinen, R. (Eds.), Towards the Sustainable Use of Europe's Forests—forest Ecosystem and Landscape Research: Scientific Challenges and Opportunities. European For. Inst., Joensuu, Finland, pp. 157–169.

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, 915–27. doi:10.1093/treephys/25.7.915

Vanclay, J.K., Skovsgaard, J.P., 1997. Evaluating forest growth models. Ecol. Modell. 98, 1–12. doi:10.1016/S0304-3800(96)01932-1

Veroustraete, F., Patyn, J., Myneni, R.B., 1994. Forcing of a simple ecosystem model with fAPAR and climatic data to estimate regional scale photosynthetic assimilation. Veg. Model. Clim. Chang. Eff. 151–177.

Vörösmarty, C.J., Sahagian, D., 2000. Anthropogenic disturbance of the terrestrial water cycle. Bioscience 50, 753–765.

Waring, R.H., Running, S.W., 1998. Forest Ecosystems: Analysis at Multiple Scales. Academic Press, San Diego.

Wellek, S., 2003. Testing Statistical Hypotheses of Equivalence. Chapman & Hall, London.

White, J.D., Running, S.W., 1994. Testing scale dependent assumptions in regional ecosystem simulations. J. Veg. Sci. 5, 687–702. doi:10.2307/3235883

White, M. a., Thornton, P.E., Running, S.W., Nemani, R.R., 2000. Parameterization and Sensitivity Analysis of the BIOME–BGC Terrestrial Ecosystem Model: Net Primary Production Controls. Earth Interact. 4, 1–85.

Wikle, C.K., Berliner, L.M., 2007. A Bayesian tutorial for data assimilation. Phys. D Nonlinear Phenom. 230, 1–16.

Williams, M., Law, B.E., Anthoni, P.M., H, M., 2001. Use of a simulation model and ecosystem flux data to examine carbon – water interactions in ponderosa pine. Tree Physiol. 287–298.

Wisskirchen, K., Tum, M., Gunther, K., Niklaus, M., Eisfelder, C., Knorr, W., 2013. Quantifying the carbon uptake by vegetation for Europe on a 1 km(2) resolution using a remote sensing driven vegetation model. Geosci. Model Dev. 6, 1623–1640.

Wu, Y., Liu, S., 2012. Automating calibration, sensitivity and uncertainty analysis of complex models using the R package Flexible Modeling Environment (FME): SWAT as an example. Environ. Model. Softw. 31, 99–109. doi:10.1016/j.envsoft.2011.11.013

Xenakis, G., Ray, D., Mencuccini, M., 2008. Sensitivity and uncertainty analysis from a coupled 3-PG and soil organic matter decomposition model. Ecol. Modell. 219, 1–16. doi:10.1016/j.ecolmodel.2008.07.020

Xiao, X., Zhang, Q., Hollinger, D., 2005. Modeling gross primary production of an evergreen needleleaf forest using MODIS and climate data. Ecol. Appl. 15, 954–969. doi:10.1890/04-0470

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L.L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A.H., Goulden, M.L., Hollinger, D.Y., Hu, Y., Law, B.E., Stoy, P.C., Vesala, T., Wofsy, S.C., 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agric. For. Meteorol. 143, 189–207. doi:10.1016/j.agrformet.2006.12.001

Zhang, X., Kondragunta, S., 2006. Estimating forest biomass in the USA using generalized allometric models and MODIS land products. Geophys. Res. Lett. 33, 1–5.

Zhu, G.F., Li, X., Su, Y.H., Zhang, K., Bai, Y., Ma, J.Z., Li, C.B., Hu, X.L., He, J.H., 2014. Simultaneously assimilating multivariate data sets into the two-source evapotranspiration model by Bayesian approach: application to spring maize in an arid region of northwestern China. Geosci. Model Dev. 7, 1467–1482. doi:10.5194/gmd-7-1467-2014

7. Tables and figures

Table 1: Characteristics of the 8 Eddy-Covariance sites considered.

Site	Calibration code	Validation code	Elevation a.s.l.	Climate	Ecosystem type	Dominant species	Coordinates (Lat, Long)
Castelporziano	Cpz07	Cpz08	9 m	Sub-humid mediterranean	Evergreen broadleaf forest	*Quercus ilex*	41.7052, 12.3761
Collelongo	Col05	Col08	1645 m	Mountain mediterranean	Deciduous broadleaf forest	*Fagussylvatica*	41.8494, 13.5881
Lavarone	Lav04	Lav06	1350 m	Mountain temperate	Evergreen needleleaf forest	*Abies alba*	45.9553, 11.2812
Nonantola	Non08	Non02	14 m	Mediterranean	Deciduous broadleaf forest	*Quercus robur*, *Fraxinus* spp.	44.6898, 11.0887
Renon	Ren05	Ren10	1794 m	Mountain temperate	Evergreen needleleaf forest	*Picea abies*	46.5878, 11.4347
Roccarespampani 2	Ro206	Ro204	165 m	Mediterranean	Deciduous broadleaf forest	*Quercus cerris*	42.3903, 11.9209
San Rossore	Sro07	Sro06	2 m	Sub-humid mediterranean	Evergreen needleleaf forest	*Pinus pinaster*	43.7279, 10.2844
Tolfa	Tol06	Tol05	415 m	Mediterranean	Mediterranean shrubland	*Arbutus unedo*	42.1897, 11.9216

Table 2: Prior probability distributions for each parameter in the Prelued model.

Parameter	Unit	Prior min.	Prior max.
β	gC mol^{-1}	0.0	1.5
γ	m^2 mol^{-1}	0.0	0.1
κ	kPa^{-1}	-10.0	0.0
X_0	°C	-100.0	0.0
τ	days	0.0	100.0
S_{max}	°C	0.0	100.0
α'	m^3 m^{-3}	0.0	1.0
ν	-	0.1	1000.0

Table 3: Model performances for all the sites considered in both calibration and validation years (R^2: coefficient of determination; RMSE: root mean square error; U: Theil's inequality coefficient; ME: modelling efficiency; ε: magnitude of region of similarity for the robust TOST test of equivalence; Dissimilarity: null hypothesis of dissimilarity of means; (C): calibration; (V): validation).

Site	Year (C)	R^2 (C)	RMSE (C)	U (C)	ME (C)	ε (C)	Dissimilarity (C)	Year (V)	R^2 (V)	RMSE (V)	U (V)	ME (V)	ε (V)	Dissimilarity (V)
Castelporziano	2007	0.69	1.09	0.2	0.67	0.25	not rejected	2008	0.49	3.45	0.51	0.42	0.25	not rejected
						0.5	rejected						0.5	not rejected
Collelongo	2005	0.96	1.04	0.14	0.96	0.25	not rejected	2008	0.5	2.60	0.44	0.35	0.25	not rejected
						0.5	rejected						0.5	not rejected
Lavarone	2004	0.84	1.45	0.2	0.82	0.25	not rejected	2006	0.84	1.33	0.21	0.84	0.25	not rejected
						0.5	not rejected						0.5	rejected
Nonantola	2008	0.93	1.27	0.18	0.92	0.25	not rejected	2002	0.84	1.96	0.31	0.82	0.25	not rejected
						0.5	rejected						0.5	not rejected
Renon	2005	0.85	1.34	0.28	0.83	0.25	not rejected	2010	0.79	2.58	0.43	0.67	0.25	not rejected
						0.5	rejected						0.5	not rejected
Roccarespampani 2	2006	0.83	2.58	0.32	0.79	0.25	not rejected	2004	0.76	2.64	0.34	0.7	0.25	not rejected
						0.5	not rejected						0.5	not rejected
San Rossore	2007	0.91	1.00	0.14	0.9	0.25	rejected	2006	0.83	1.15	0.17	0.83	0.25	not rejected
						0.5	rejected						0.5	rejected
Tolfa	2006	0.37	2.92	0.38	0.32	0.25	not rejected	2005	0.02	2.78	0.4	-0.18	0.25	not rejected
						0.5	not rejected						0.5	not rejected

Table 4: Derived and modeled GPP for all the sites considered, with uncertainties associated to the average daily value of GPP ($gC \, m^{-2} \, d^{-1}$). Only the days of simulation were both data and model results were present were used. Uncertainties were calculated as the average of the lowest (min) and highest (max) values of daily uncertainties. (C): calibration; (V): validation.

Site	Year (C)	N° data (C)	Average daily GPP derived (min-max) (C)	Average daily GPP modeled (min-max) (C)	Year (V)	N° data(V)	Average daily GPP derived (min-max)(V)	Average daily GPP modeled (min-max)(V)
Castelporziano	2007	297	5.01 (4.23-5.79)	4.71 (4.53-4.86)	2008	231	5.27 (4.33-6.21)	3.95 (3.75-4.22)
Collelongo	2005	281	5.03 (4.05-6.01)	4.76 (4.58-4.99)	2008	335	5.17 (4.34-6.00)	6.64 (6.49-7.14)
Lavarone	2004	209	6.27 (5.27-7.28)	5.77 (5.55-6.02)	2006	323	5.27 (4.38-6.15)	5.36 (5.12-5.60)
Nonantola	2008	358	5.21 (4.29-6.13)	4.81 (4.59-4.98)	2002	244	4.44 (3.58-5.31)	4.25 (3.99-4.52)
Renon	2006	330	3.57 (2.84-4.30)	3.11 (2.98-3.25)	2010	349	4.19 (3.34-5.03)	2.76 (2.66-2.88)
Roccarespampani 2	2006	335	5.95 (4.87-7.03)	5.15 (5.01-5.27)	2004	280	6.35 (5.30-7.40)	5.05 (4.93-5.17)
San Rossore	2007	333	6.61 (5.58-7.64)	6.35 (6.13-6.57)	2006	356	6.06 (5.11-7.00)	6.28 (6.03-6.52)
Tolfa	2006	280	6.94 (5.88-8.00)	5.59 (5.36-5.82)	2005	223	6.49 (5.50-7.47)	5.77 (5.42-5.99)

Table 5: Calibrated parameter values (as the parameter vector with the maximum log-likelihood) and their uncertainties (as the standard error of the posterior distribution for each parameter, excluding the burn-in phase).

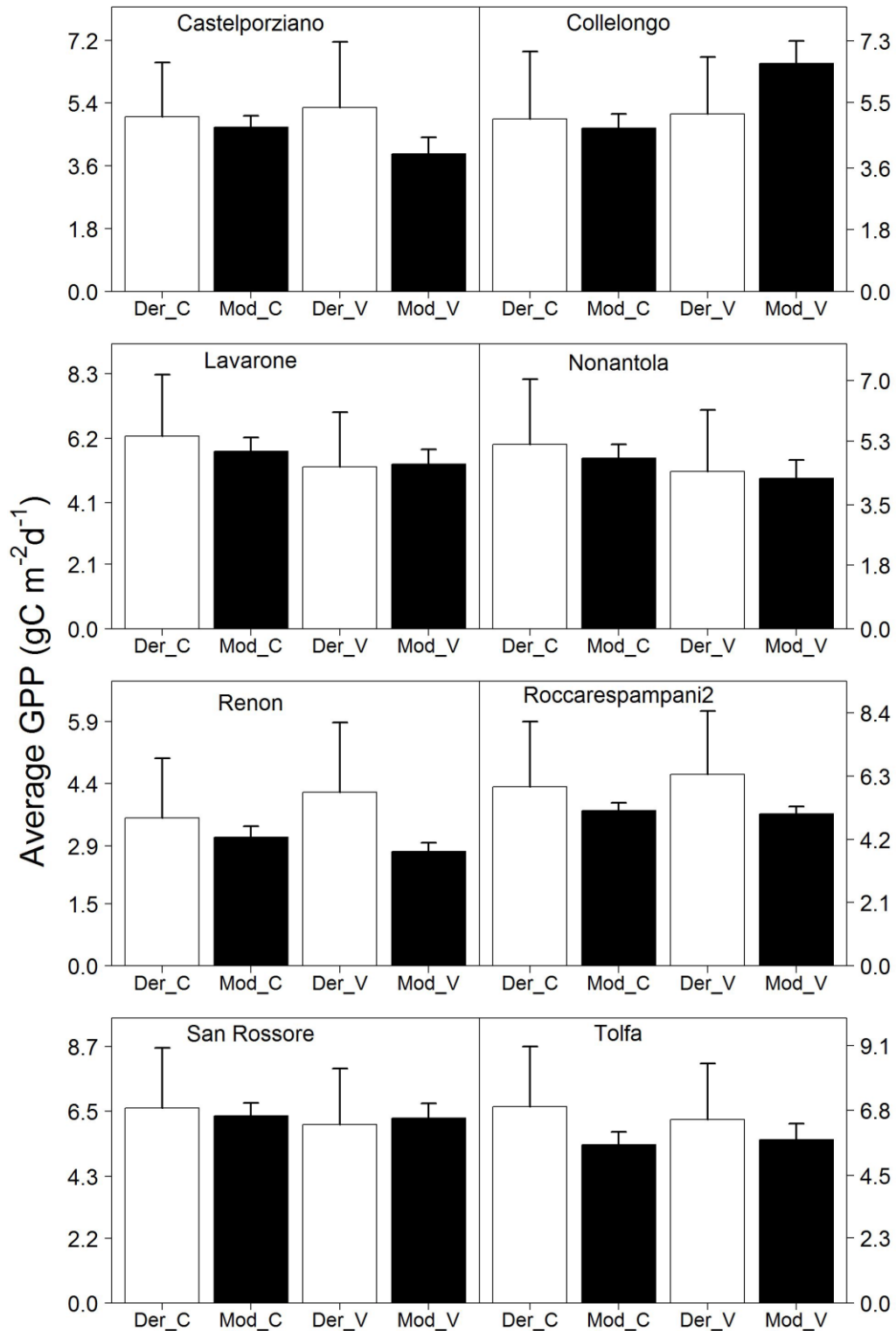|	Castelporziano	Collelongo	Lavarone	Nonantola	Renon	Roccarespampani 2	San Rossore	Tolfa
β	0.888	0.364	0.605	1.495	0.507	0.817	0.498	1.309
(se)	(0.080)	(0.006)	(0.012)	(0.018)	(0.008)	(0.074)	(0.006)	(0.031)
γ	0.039	0.010	0.023	0.021	0.031	0.000	0.022	0.084
(se)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.003)
κ	-0.568	-0.209	-0.918	-0.655	-0.555	0.000	-0.118	-0.244
(se)	(0.017)	(0.026)	(0.026)	(0.020)	(0.032)	(0.002)	(0.020)	(0.021)
X_0	-27.280	-0.008	-8.909	-0.620	-2.707	-0.003	-0.120	-0.007
(se)	(2.448)	(0.030)	(0.229)	(0.200)	(0.157)	(0.018)	(0.179)	(0.031)
τ	0.965	91.362	6.420	3.065	7.466	4.513	7.154	4.309
(se)	(0.058)	(2.800)	(0.265)	(0.145)	(0.186)	(0.276)	(1.170)	(0.208)
S_{max}	71.539	3.284	17.642	28.658	8.967	80.233	10.097	9.235
(se)	(5.434)	(0.212)	(0.522)	(0.621)	(0.285)	(7.192)	(0.296)	(0.146)
α'	0.051	0.393	0.430	0.356	0.225	0.453	0.034	0.227
(se)	(0.002)	(0.091)	(0.090)	(0.005)	(0.099)	(0.089)	($2 \cdot 10^{-4}$)	(0.016)
ν	1.473	0.787	802.403	2.517	706.691	889.731	4.801	0.558
(se)	(0.066)	(114.435)	(100.403)	(0.149)	(102.518)	(98.658)	(0.133)	(0.037)

Figure 1.Barplot of modeled and derived average daily GPP for each site considered (Mod_C: modeled GPP in the calibration year, Mod_V: modeled GPP in the validation year, Der_C: derived GPP in the calibration year, Der_V: derived GPP in the calibration year), with their uncertainties (error bars). Only the days of simulation where both data and model results were present were used. Uncertainties were calculated as the difference of the averages of the lowest and highest values of daily uncertainties.
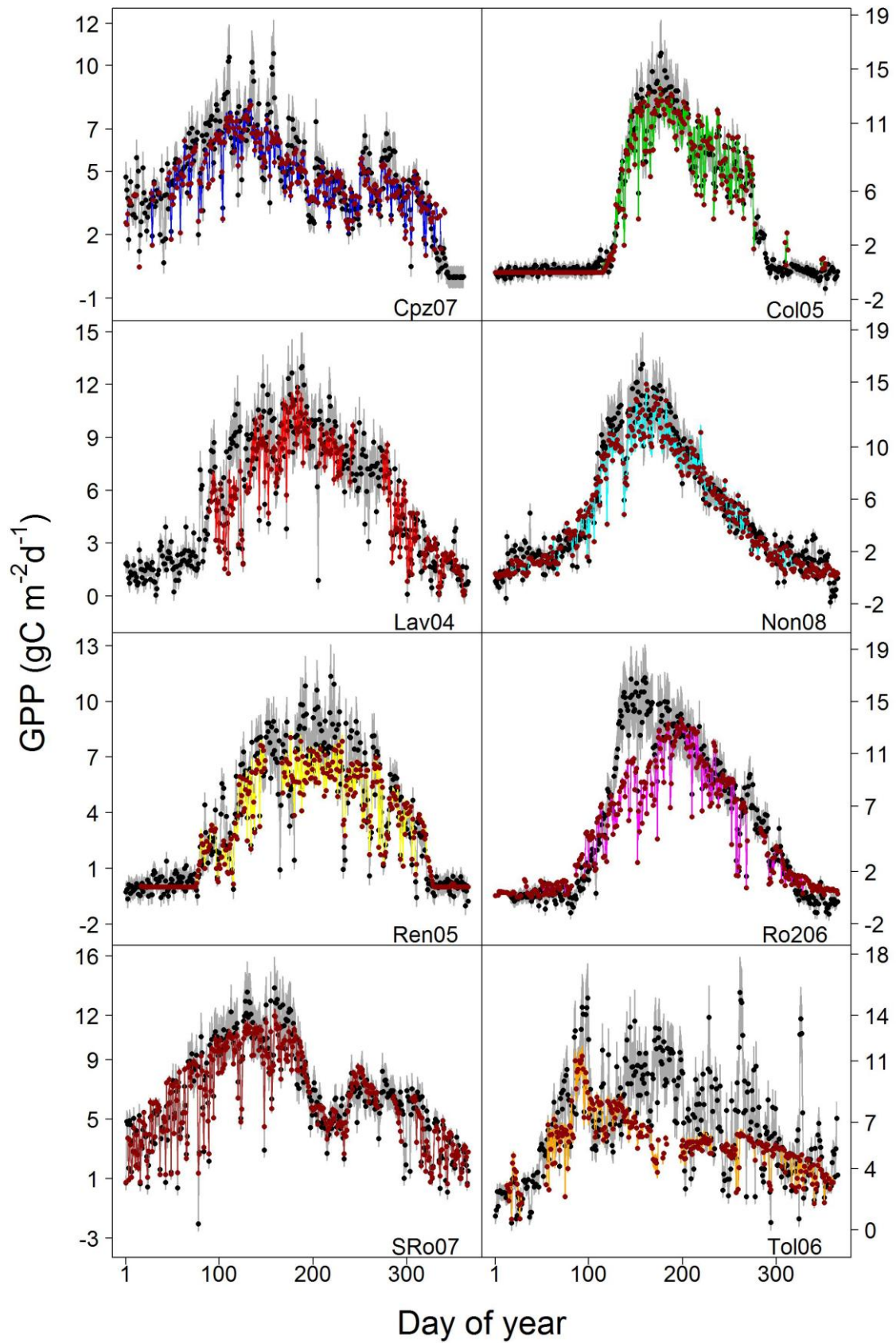
Figure 2. Derived (black dots and grey polygons) and modeled (red dots and color polygons) daily GPP and uncertainties for the sites considered in the calibration year. Site codes are listed in table 1.
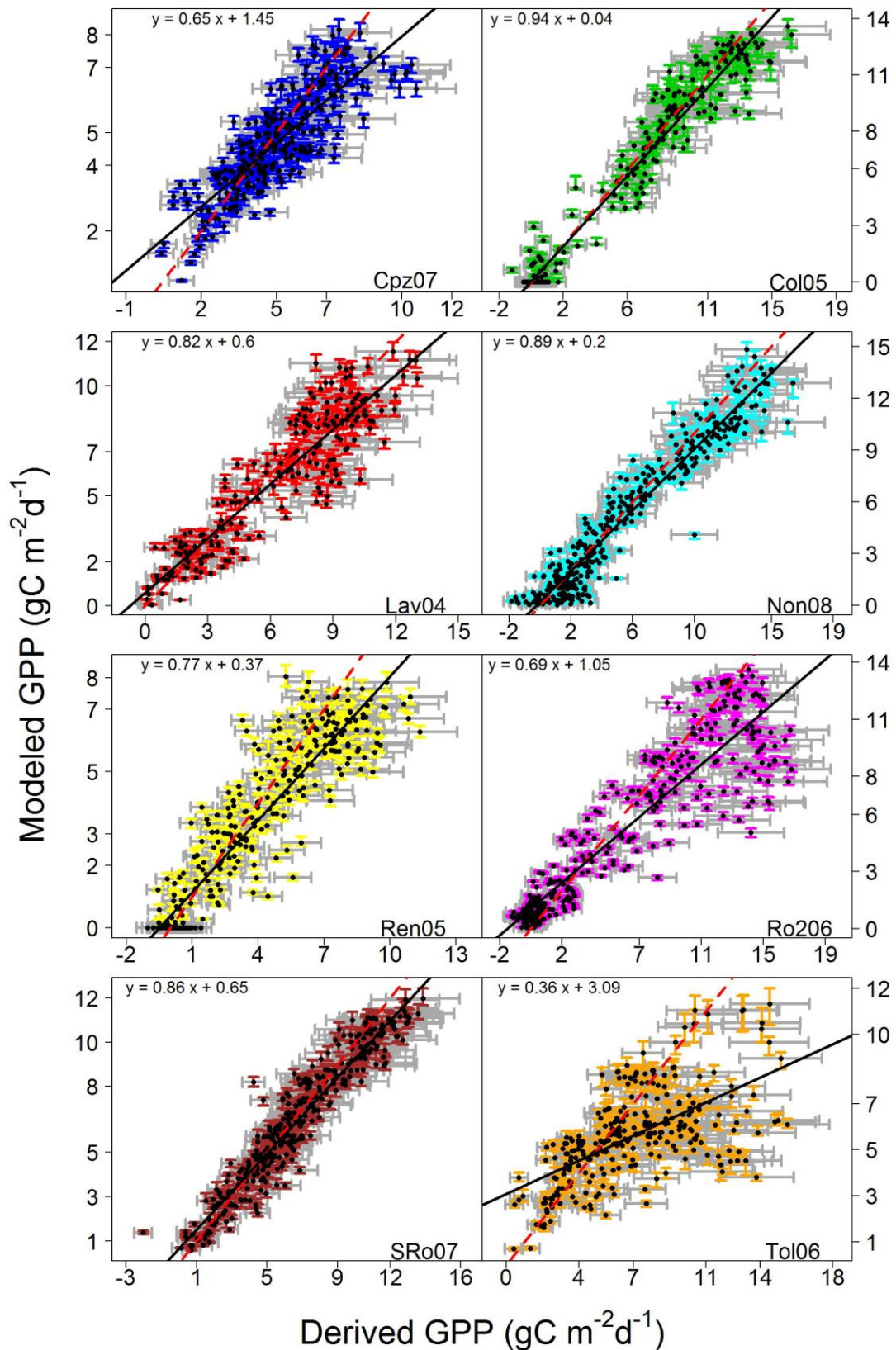
Figure 3. Goodness-of-fit between derived and measured GPP for the sites considered in the calibration year (black dots: daily GPP, grey bars: data uncertainties, color bars: model uncertainties, black line: regression line, red dotted line: 1:1 line).Site codes are listed in table 1.
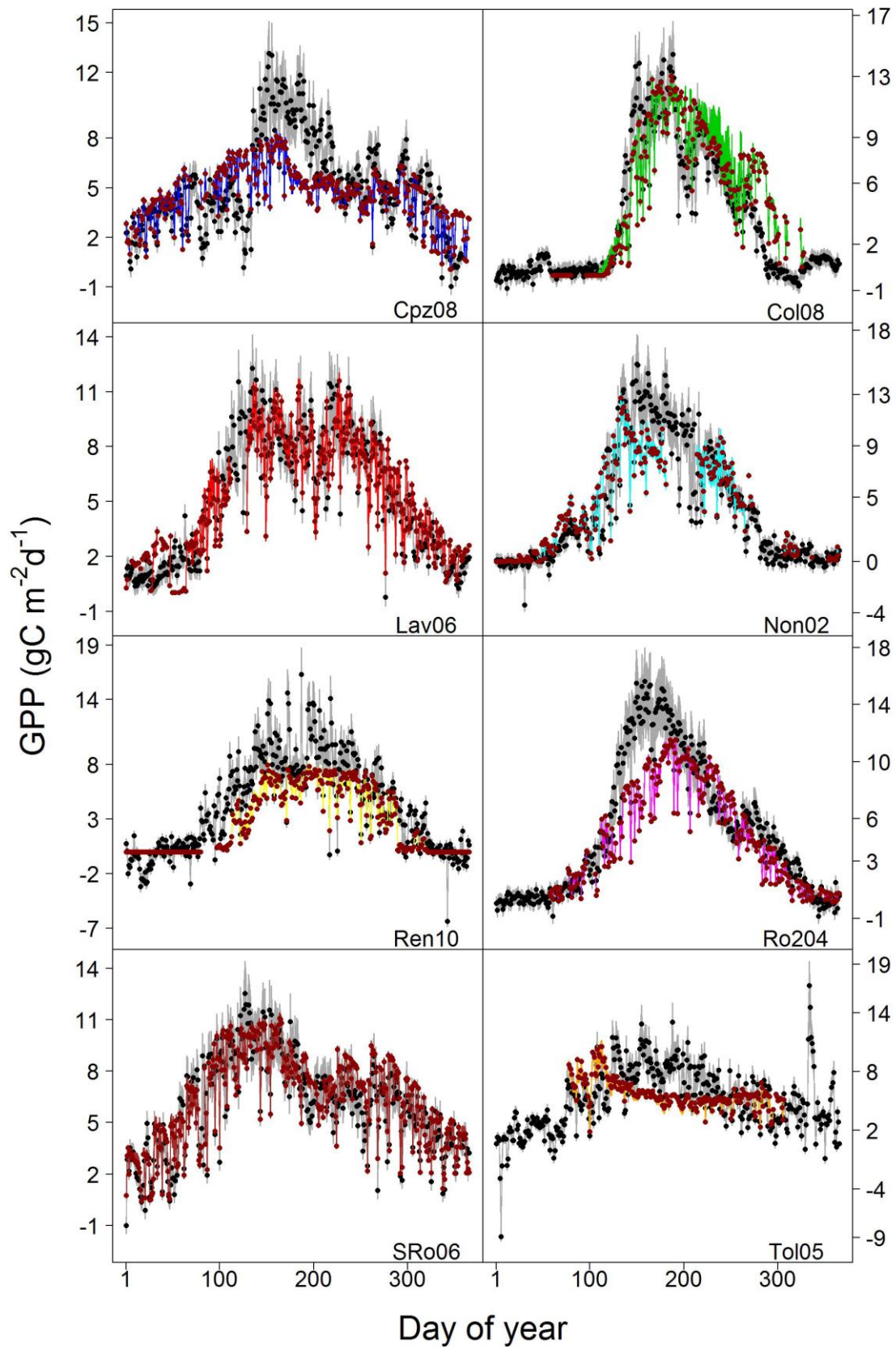
Figure 4. Derived (black dots and grey polygons) and modeled (red dots and color polygons) daily GPP and uncertainties for the sites considered in the validation year. Site codes are listed in table 1.
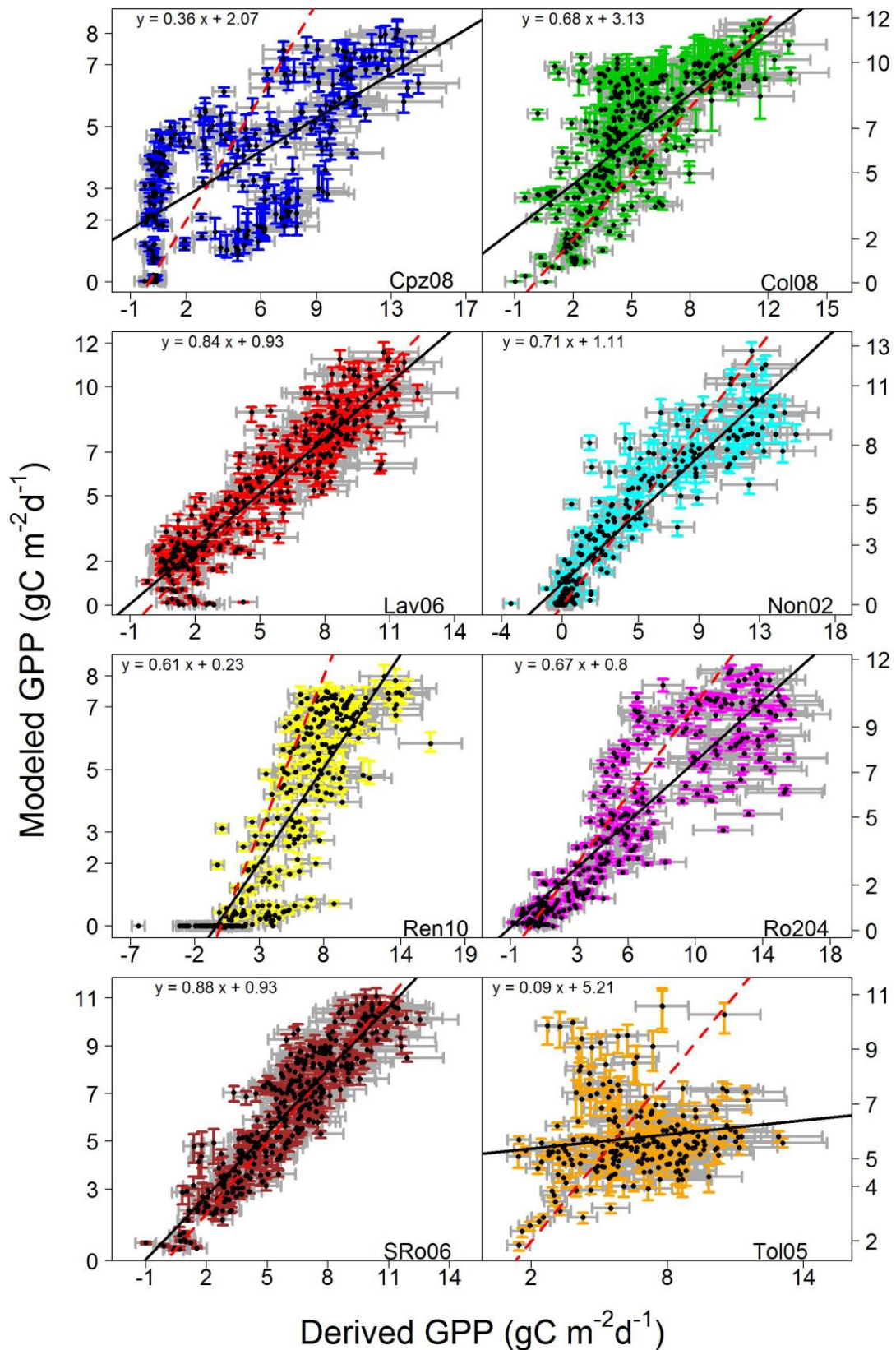
Figure 5. Goodness-of-fit between derived and measured GPP for the sites considered in the validation year (black dots: daily GPP, grey bars: data uncertainties, color bars: model uncertainties, black line: regression line, red dotted line: 1:1 line).Site codes are listed in table 1.
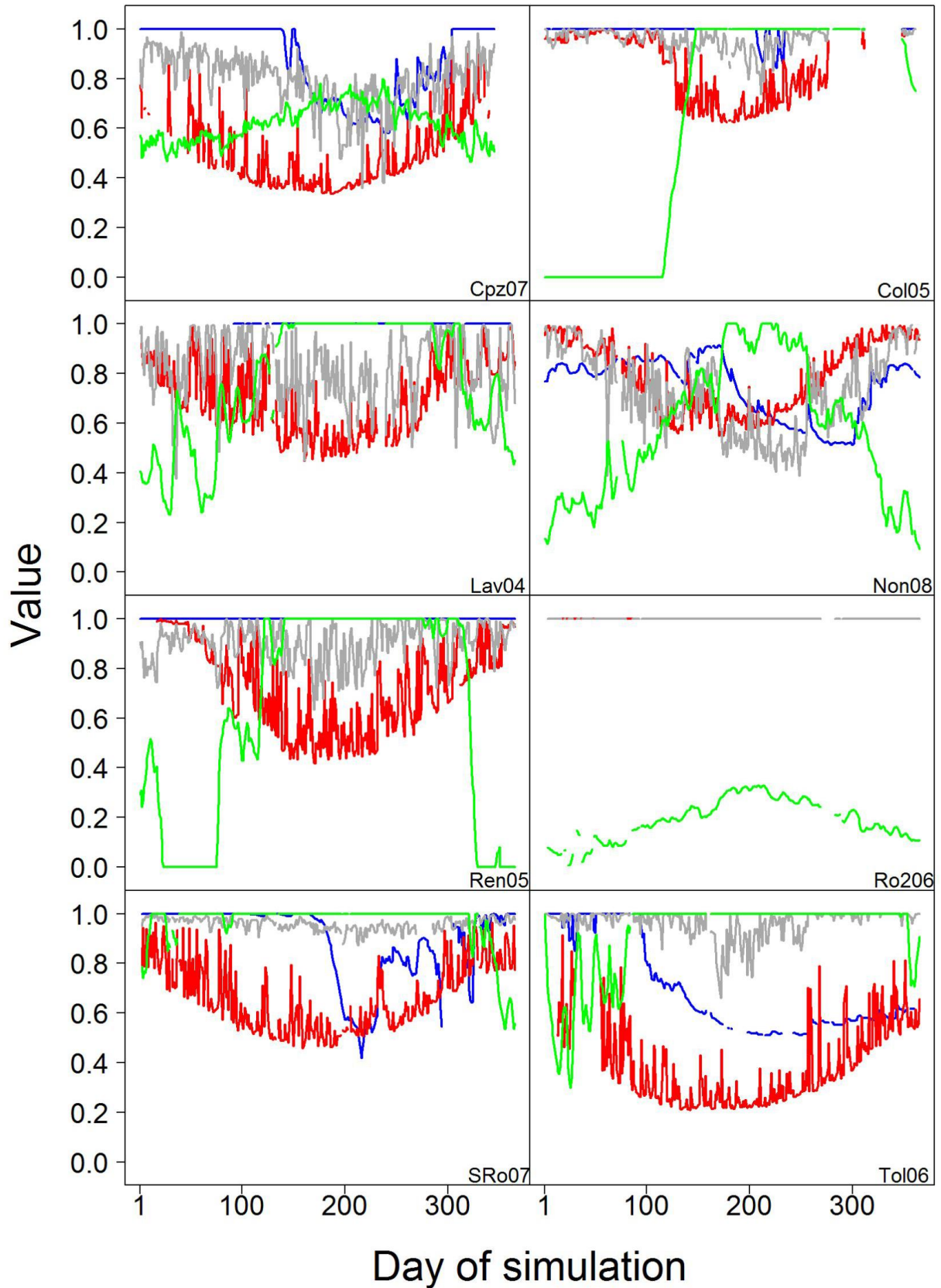
Figure 6. Daily values of modifiers in the calibration year for all the sites considered (green: temperature modifier, blue: SWC modifier, grey: VPD modifier, red: light modifier). Site codes are listed in table 1.

54

Chapter 3: A user-friendly forest model with a multiplicative mathematical structure: a Bayesian approach to calibration

Abstract

Forest models are being increasingly used to study ecosystem functioning, through the reproduction of carbon fluxes and productivity in very different forests all over the world. Over the last two decades, the need for simple and "easy to use" models for practical applications, characterized by few parameters and equations, has become clear, and some have been developed for this purpose. These models aim to represent the main drivers underlying forest ecosystem processes while being applicable to the widest possible range of forest ecosystems. Recently, it has also become clear that model performance should not be assessed only in terms of accuracy of estimations and predictions, but also in terms of estimates of model uncertainties. Therefore, the Bayesian approach has increasingly been applied to calibrate forest models, with the aim of estimating the uncertainty of their results, and of comparing their performances.

Some forest models, considered to be user-friendly, rely on a multiplicative or quasi-multiplicative mathematical structure, which is known to cause problems during the calibration process, mainly due to high correlations between parameters. In a Bayesian framework using a Markov Chain Monte Carlo sampling this is likely to impair the reaching of a proper convergence of the chains and the sampling from the correct posterior distribution. Here we show two methods to reach proper convergence when using a forest model with a multiplicative structure, applying different algorithms with different number of iterations during the Markov Chain Monte Carlo or a two-steps calibration. The results showed that recently proposed algorithms for adaptive calibration do not confer a clear advantage over the Metropolis-Hastings Random Walk algorithm for the forest model used here. Moreover, the calibration remains time consuming and mathematically difficult, so advantages of using a

fast and user-friendly model can be lost due to the calibration process that is needed to obtain reliable results.

1. Introduction

Gross Primary Production (GPP) is a key component of the carbon balance. Therefore, it is the central output of many forest ecosystem models (De Weirdt et al., 2012; Mäkelä et al., 2000b; Tjiputra et al., 2013), and is being increasingly targeted by remote sensing applications as a proxy to predict global carbon fluxes and plant light-use efficiency at large spatial scales (Still et al., 2004; Wisskirchen et al., 2013; Zhang and Kondragunta, 2006). It can also be estimated by the Eddy-covariance technique: this micrometeorological method computes the net CO_2 turbulent flux between a given ecosystem and the atmosphere from the covariance between the fluctuations of vertical wind velocity and CO_2 concentrations, averaged at a half-hour time scale. This can be used to estimate both ecosystem respiration (Re) and GPP. Currently, a global network of more than 500 EC stations exist worldwide to continuously monitor the CO_2 and energy exchange between ecosystems and the atmosphere, whose homogeneity is ensured by similar standardized procedures(Baldocchi, 2008). Despite extensive efforts and several techniques tested, GPP quantification remains challenging in most ecosystems given its dependence on several meteorological, environmental and internal drivers at several time scales.

Most of the models of forest growth and biogeochemical cycles are detailed, multi-variable models that need much environmental information and careful parameterisation before they can be run (Landsberg & Waring 1997). To make them suitable for a wider range of purposes and accessible to a wider range of users, a process of simplification started in the 90's (White & Running 1994; Landsberg &Waring 1997) with the aim of developing models that could be of use in applied forest management. One step in this direction was represented by the creation of hybrid models (e.g. FORCYTE-11 (Kimmins, 1986)), combining the

predictive power of process-based models with the short-term believability of mensuration-based models (Kimmins et al., 1999; Medlyn et al., 2003). Unlike full process-based models, hybrid models are based on the principle that only the processes that are expected to change would be included in the modelling effort, reducing the number of processes taken into account and resulting in a simplification of the overall model structure (Kimmins et al., 2008).

A widely used group of simple models for GPP is based on the concept of Light Use Efficiency (LUE), defined as the ratio of GPP to Absorbed Photosynthetically Active Radiation. They mainly rely on simplified physiological processes and empirical parameters, require little information to be run, the computations are usually fast, and their mathematical structure is often quasi- or totally multiplicative. These models assume that vegetation has a potential LUE, which can be described as the ability of plants to use light for photosynthesis in absence of limiting factors, decreased by modifying factors that account for suboptimal conditions for photosynthesis(Landsberg and Waring, 1997; McMurtrie et al., 1994).

Some examples of these models are 3PG (Landsberg & Waring 1997), C-Fix (Veroustraete et al., 1994), the model developed by Horn and Schulz(2011b), and Prelued (Mäkelä et al., 2008).

Despite relying on a multiplicative mathematical structure and on several empirical parameters, of which little is known in the literature, Prelued has been successfully applied to several ecosystems all across the world (Mäkelä et al., 2008; Peltoniemi et al., 2012b). Compared to the majority of the LUE-based models that work at monthly or annual time scale, relying on a linear relationship between GPP and Absorbed Photosynthetically Active Radiation (APAR) and on a parabolic effect of temperature, the Prelued model calculates GPP at a daily time scale, basing the calculations on a nonlinear relationship between APAR and GPP (Medlyn et al., 2003;Turner et al., 2003), a saturating effect of average daily temperature(which simulates the ecosystem "acclimation" to temperature, Mäkelä et al. (2004)), and daily meteorological and environmental variables. The importance of these

environmental variables has been recently highlighted by McCallum et al. (2013), applying the Prelued model at four Eddy-Covariance sites in Russia: this clearly demonstrates the improved fit of the model when considering both temperature acclimation and nonlinearity in response to APAR, especially in temperature-controlled ecosystems.

The Bayesian approach to model calibration has become more and more popular in the last few years to obtain insights on both model predictions and uncertainties. The main characteristic of a Bayesian calibration is that it quantifies model inputs and outputs in the form of probability distributions, and applies the rules of probability theory to update the distributions when new data are obtained (Sivia, 1996; van Oijen et al., 2005). This approach has been widely used in the past in different fields, and recently it has been applied to different kind of ecosystem models, focusing on both crop (Zhu et al., 2014) and forest (van Oijen et al., 2005; Svensson et al., 2008; Chevallier et al., 2006; van Oijen et al., 2011; Van Oijen et al. 2013). Even so, the application of the Bayesian method to LUE-based models is not as common as its application to process-based models, with a very few studies heading in this direction (Still et al., 2004; Xenakis et al., 2008).

Implementations of Bayesian calibration rely mainly on Markov Chain Monte Carlo (MCMC) algorithms for sampling the parameter space to obtain posterior probability distributions for the model parameters. The efficiency of the MCMC technique is highly dependent on the model structure. The high correlations between parameters that are induced by a multiplicative model structure generally make the convergence of the MCMC more difficult, impairing the reliability of the results of the calibration. Different methods have been implemented to avoid or reduce such problems: the use of very long chains (Gilks et al., 1996), model re-parameterization to avoid strong correlations (Buzzi-Ferraris & Manenti 2010; Gilks et al. 1996), and the use of more efficient algorithms (Gilks et al., 1996; ter Braak, 2006).In this study, three algorithms characterized by increasing complexity and

efficiency were applied: the Metropolis-Hastings Random Walk (MHRW), the Adaptive Metropolis (AM), and the Differential Evolution Markov Chain (DEMC).

The Metropolis-Hastings random walk (MHRW) (Casella and Robert, 1999) algorithm produces a walk through the parameter space such that that the collection of visited points forms the desired sample from the posterior distribution, discarding some initial values (van Oijen et al.; 2005).At each iteration of the algorithm, a new candidate parameter vector is proposed stochastically, i.e. the jump from the current point to the proposed next one follows a probability distribution. The most commonly used proposal distribution is the multivariate Gaussian. Whether the proposal is accepted, depends on the prior probabilities and likelihoods of the current and proposed parameter vectors. In the MHRW, the proposal distribution itself does not change, so average proposed jump directions and distances remain the same throughout the random walk. This is different in the next two MCMC algorithms. The adaptive Metropolis algorithm (AM) is a modification of the MHRW. The key attribute of the AM algorithm is the continuous adaptation of its proposal distribution. The adaptation consists of gradual convergence of the covariance matrix of the proposal distribution to the covariance matrix of the parameters visited so far in the chain (Haario et al. 2001; Smith & Marshall 2008). The differential evolution Markov chain algorithm (DEMC) is formed by combining the differential evolution algorithm of Storn & Price (1997), designed for global optimization in real parameter spaces, with MCMC sampling, utilizing standard Metropolis principles. The result is a population MCMC algorithm, where multiple chains are run in parallel and allowed to learn from each other. This combination intends to overcome the difficulties common to MCMC methods of choosing an appropriate scale and orientation (respectively the size of each jump in the MCMC sampling and its direction in the parameter space) for the proposal distribution, while also addressing issues of computational efficiency related to the time to reach convergence(Smith and Marshall, 2008; ter Braak, 2006).

This work aims at testing different procedures that could be successfully applied to the variety of forest models with similar structure to reach proper convergence during the MCMC sampling. We applied a Bayesian calibration with different algorithms and number of iterations, as well as reparameterization and multiple steps calibration, to the Prelued model, employed as a case study, using one year of daily GPP data from an Eddy-Covariance (EC) tower in the Italian Alps.

2. Materials and Methods

2.1 Model formulation

The model used as a case study is a modified version of a LUE-type model of daily photosynthetic production of the canopy (Mäkelä et al., 2008):

$$GPP_j = \beta APAR_j \prod_i F_{ij} \qquad (1)$$

where GPP_j is canopy Gross Primary Production (gC m^{-2}) during day j, β is potential Light Use Efficiency (gC mol^{-1}), $APAR_j$ is Absorbed Photosynthetically Active Radiation (mol m^{-2}) during day j, and $F_{ij} \in [0, 1]$ are modifying factors accounting for suboptimal conditions on day j. The actual LUE of the canopy on day j is the product of β and the current values of the modifiers.

To account for the nonlinearity in the response to APAR, a light modifier FL was defined so as to yield the rectangular hyperbola when multiplied with the linear response included in the LUE model:

$$FL_j = 1/(\gamma APAR_j + 1) \qquad (2)$$

where γ (m^2 mol^{-1}) is an empirical parameter. The effect of temperature on daily GPP was modeled using the concept of state of acclimation, S_j (°C) (Mäkelä et al., 2004), a piecewise linear function of X_j (°C) calculated from the mean daily ambient temperature, T_j (°C), using a first-order dynamic delay model:

$$X_j = X_{j-1} + (1/\tau) (T_j - X_{j-1}), X_1 = T_1 \qquad (3)$$

$$S_j = \max \{X_j - X_0, 0\} \qquad (4)$$

where τ (days) is the time constant of the delay process and X_0 (°C) is a threshold value of the delayed temperature. The modifying function FS is defined as

$$FS_j = \min \{S_j/ S_{max}, 1\} \qquad (5)$$

where the empirical parameter S_{max} (°C) determines the value of S_j at which the temperature modifier attains its saturating level.

Following Landsberg & Waring (1997) the Vapour Pressure Deficit (VPD) modifier was defined as

$$FD_j = e^{\kappa VPDj} \qquad (6)$$

where VPD_j (kPa) is VPD in day j and κ (kPa^{-1}) is an empirical parameter assuming typically negative values.

2.2 Data

The data for the Italian Eddy Covariance site of Lavarone for the year 2004 has been downloaded from the European Fluxes Database Cluster (www.europe-fluxdata.eu). Lavarone is a ca. 130 years old alpine coniferous forest, dominated by Silver fir (*Abies alba* Mill.) and Norway spruce (*Picea abies* (L.) Karst.), with minor presence of European beech (*Fagus sylvatica* L.) and located at 1350 m a.s.l. in the Trento province, western Italian Alps. The site characteristics of Lavarone are described in detail in Rodeghiero & Cescatti (2005).

Daily air temperature, relative humidity (Rh) and PAR were used as input data. Daily VPD was calculated from Rh and air temperature following Allen et al. (1998), while daily GPP was used to calibrate the model. Daily APAR was calculated following Mäkelä et al.(2008), using Normalized Difference Vegetation Index (NDVI) data as a proxy for fAPAR: for that purpose, NDVI data with 0.25 km spatial grid and 16 days time-step were downloaded from the MODIS repository (http://daac.ornl.gov/cgi-bin/MODIS/GLBVIZ_1_Glb/modis_subset_order_global_col5.pl).

Missing data for either a weather variable or GPP resulted in a missing outcome of the model for that day *j*. Therefore, 292 data points were actually used to calibrate the model.

The Bayesian calibration requires an estimate of the uncertainties around the data used during the calibration(van Oijen et al., 2005). Uncertainties around GPP_j were calculated as follows:

$$GPP_j = GPP_j \pm y_j \qquad (7)$$

where y_j is the maximum of 1 gC m-2 and a random value in the interval $[1:0.3*GPP_j]$.The lower bound of 1 g C m^{-2} is necessary to ensure that low values of GPP_j would not get an overwhelming weight during the calibration procedure.

2.3 Bayesian calibration

The prior parameter distributions were set based on the information made available by (Mäkelä et al., 2008) and Peltoniemi et al. (2012). Since the parameter distributions were partly unknown, and since many parameters are empirical and without physiological meaning, we set the prior distributions as uniform distributions (i.e. any value has the same probability to occur) and wide enough to cover a very wide range of possible values (table 1).

To investigate in detail the model behavior during a Bayesian calibration and to tackle the issues related to slow convergence, we tested four different procedures:

1. *Single-step calibration: f*or each of the three algorithms applied (MHRW, AM, DEMC) different simulations with an increasing number of iterations were performed to test the efficiency of each algorithm in reaching convergence. Three simulations were run, with 10^4, 10^5 and 10^6 iterations in total for each algorithm. An initial burn-in phase was set to 30% of the total number of iterations for all the algorithms. For the DEMC algorithm, 100 chains were considered, making the number of iterations per chain respectively 10^2, 10^3 and 10^4.

2. *Model comparison:* we ran a second LUE-based model with slightly different structure, on the same data and with the same calibration settings. The model chosen for this purpose was the model developed by Horn and Schulz(2011a)as described in Horn and Schulz(2011b). An initial burn-in phase was set to 30% of the total number

of iterations for all the algorithms. For the DEMC algorithm, 100 chains were considered, making the number of iterations per chain respectively 10^2, 10^3 and 10^4.The main difference in the mathematical structure between the two models is that while in Prelued GPP is calculated as a product of potential LUE, APAR, and modifiers (Eq. 1), in Horn and Schulz(2011b)GPP is calculated as:

$$GPP_j = LUE\ APAR_j[\ pFT_j + (1\text{-}p)\ FW_j] \qquad (8)$$

with GPP_j (gC m^{-2}) denoting the gross flux of carbon uptake in day j, LUE(gC MJ^{-1}) being the maximum attained Light Use Efficiency, APAR (MJ m^{-2}) the Absorbed Photosynthetically Active Radiation in day j, and p a weighting factor for the modifiers FT and FW.

FT is a sigmoidal peak function defined as:

$$FT = 4\ e^{\ -(Ts\text{-}Topt)/kT} / (1+e^{-(Ts\text{-}Topt)/kT})^2 \qquad (9)$$

where Ts is the soil temperature (°C), $Topt$ (°C) is the temperature at which the light use efficiency is maximum, and kT (°C^{-1}) is the rate of change from the lower level of FT to its maximum.

FW is defined as following sigmoidal function:

$$FW = 1\ /\ (1+e^{kW(W\text{-}Wi)}) \qquad (10)$$

where W is a moisture surrogate (in our case the Soil Water Content (m^3 m^{-3})),kW is the constant rate of change between lower and upper level (set to -13.1 following Horn and Schulz(2011b))and Wi is the inflection point with units depending on the choice of W.

Following Jarvis et al.(2004), a lag function was applied to Ts:

$$ZF_j = (1\text{-}\alpha)\ Ts_j + \alpha ZF_{j\text{-}1} \qquad (11)$$

where α (-) is the lag parameter. Eq. (11) is only applied to Ts, considered the dominant driver of the vegetation stands; this main driver is expected to trigger the

start and end of dormant periods after which the vegetation has to regenerate and redevelop green tissue (J. E. Horn and Schulz, 2011).

FT and FW are scaled between 0 and 1 and describe the dependence of the Light Use Efficiency on the soil temperature and a moisture surrogate. Overall, the structure of this latter model is less multiplicative than Prelued, which should make its calibration easier. The prior distributions for this model have been derived from tab. 2 in Horn and Schulz (2011a), using the minimum and maximum value for each parameter as boundaries and keeping the distributions uniform.

3. *Reparameterization:* we reformulated four parameters of Prelued out of six, changing their meaning and the model formulation accordingly:

$$\beta' = \beta/\gamma \qquad (12)$$

$$\gamma' = 1/\gamma \qquad (13)$$

$$Smax' = 1/Smax \qquad (14)$$

$$X0' = X0/Smax \qquad (15)$$

Given the purpose of this approach was reaching faster the convergence only two calibrations instead of three (10^4 and 10^5 iterations) were performed.

4. *Two-steps calibration:* in this procedure, the posterior correlations between parameters found in the first step were used to reduce the number of parameters involved in the second step. If two parameters were strongly correlated, the coefficients of the linear regression between them were used to estimate one parameter as a function of the second, reducing the number of calibrated parameters. These coefficients were calculated for each number of iterations, and used in a second calibration with the same length: the coefficients calculated after the 10^4 iterations calibration were used for a second 10^4 iterations calibration with reduced number of parameters, and the same approach was used for the 10^5 iterations calibration. Since convergence was

reached after the first step for the longest calibration, only two second steps were run (10^4 and 10^5 iterations).

2.4 Measure of convergence

The reaching of the convergence region was visually assessed, along with the different behavior of the Markov Chain between different numbers of iterations and their similar behavior between algorithms. In order to obtain a quantitative measure of convergence of the chains to the posterior distribution, the last 50% of the longest chain for each algorithm were split in half and the means and variances of the two halves were compared. For the DEMC algorithm, only the chain with maximum log-likelihood was chosen for this purpose.

3. Results

3.1 Bayesian calibration

3.1.1 Single-step calibration

For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 iterations, approached convergence at 10^5 iterations, and reached good convergence at 10^6 iterations (figure 1).The same pattern emerged from the analysis of the posterior distributions: for many parameters, the posterior distributions were bimodal, shifted, or as broad as the priors at 10^4 iterations, while becoming leptokurtic at 10^6 iterations for all the parameters. With the latter number of iterations, the posterior distribution thus narrowed the parameter space, converging in the same region (figure 2). The parameter sets with best log-likelihood (table 2) and the posterior correlation coefficients between parameters (table 3) were very similar between algorithms with only few exceptions. This confirmed the convergence on the same joint posterior distribution and not only on the marginal distributions for each parameter. The strongest correlation was found for the threshold value of the delayed temperature X_0 and the empirical parameter S_{max} (correlation coefficient varies from -0.923 to -0.928 depending on the algorithm), both involved in the response to temperature. Strong correlation existed also

between the Potential LUE β, and the empirical parameter γ (correlation coefficient varies from 0.89 to 0.91 depending on the algorithm), which were both involved in the response to APAR. Concerning the log-likelihood values of the best parameter set, the MHRW algorithm showed the best result compared to the AM and the DEMC (table 2).

3.1.2 Model comparison

The application of the less multiplicative LUE-based model developed by Horn and Schulz(2011b) to the same dataset did not show better results compared to Prelued, in terms of reaching proper convergence, even at a high number of iterations. For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 and 10^5 iterations, and reached convergence at 10^6 iterations for some parameters only (figure 3). The analysis of the posterior distributions showed the same trends as in Prelued: for many parameters, the posterior distributions were bimodal, shifted, or as broad as the priors at 10^4 iterations, while narrowing the parameter space at 10^6iterations and converging in the same region (figure 4).

3.1.3 Reparameterization

The alternative formulation proposed to overcome the convergence problems with the calibration of Prelued did not result in faster convergence. For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 and 10^5 iterations (figure 5). The analysis of the posterior distributions showed a situation far from convergence in every case: for many parameters, the posterior distributions were bimodal, shifted, or as broad as the priors at both 10^4 and10^5 iterations, sometimes exploring different regions of the parameter space (figure 6).

3.1.4 Two-steps calibration

The posterior correlations found in the first step of calibration described in paragraph 3.1 (parameters β and γ, and parameters τ and S_{max}) were used to reduce the number of parameters estimated in the second step. In particular, γ was estimated as a linear function of β, and S_{max}

as a linear function of τ. For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 iterations, but reached good convergence at 10^5 iterations for all the parameters (figure 7). The analysis of the posterior distributions showed well defined distributions at both 10^4 and 10^5 iterations, and in the latter case they converged in the same region of the parameter space (figure 8).

3.2 *Quantitative measure of convergence*

For the single-step calibration with 10^6 iterations the means of the first- and second-halves of the MCMC were within 1.5% of each other except for parameter τ (time constant of the delay process in response to temperature) using the DEMC algorithm (2.9 %). The variances of the first- and second-halves were within 15% except for parameter X_0 (i.e., the threshold value of the delayed temperature) using the DEMC algorithm (23.2 %) (table 4). This proves that the longest chains for each parameter and algorithm are converging on the same region of parameter space (figure 9).

4. Discussion

In disagreement with the expectation, given their increasing complexity and efficiency, all three MCMC-methods tested in this study were similarly effective. Although this similarity in behavior between algorithms was a surprising result, the main output of this study was that a very high number of iterations was required for each of the three calibration algorithms to stabilize in the convergence region. This is especially true considering the simplicity of the Prelued model. This 6-parameters empirical model required 10^6 iterations to reach convergence, whereas a 39-parameter mechanistic forest model was calibrated with chains of length 10^5 (van Oijen et al., 2005), and 10^5 iterations were enough to allow proper convergence for 4 process-based models with higher complexity than Prelued (Marcel van Oijen et al., 2011). Despite the high number of iterations required, all three algorithms reached convergence since they all explored the same parameter space (figure 3). This

demonstrates that the three algorithms were not attracted in different regions of the parameter space, which excludes the risk of undiagnosed slow convergence (Gilks et al., 1996).

The LUE model by Horn and Schulz(2011b) showed the same convergence problems as Prelued when calibrated with a Bayesian approach, despite a less multiplicative structure (figure 3). Therefore, the comparison of these two models suggested that the multiplicative structure of Prelued was probably the main factor responsible for the difficulties in the calibration, but is unlikely to be the only one. It should also be considered that even if LUE-type models are largely empirical, they usually rely on parameters with physiological meaning. Their analysis thus gives insights about the ecosystem characteristics and behavior, and allows for comparison between different models. For example, the well-known and widely applied 3PG model (Landsberg &Waring 1997) has the same mathematical properties of Prelued, even if not so extreme. Therefore, beside the strong multiplicative mathematical structure, the problems in calibrating Prelued were likely due to the indefinite nature of the empirical parameters, neither ecological nor physiological. This renders the prior distributions difficult to specify and forces the MCMC to investigate a broad parameter space, delaying the identification of the convergence region.

The reparameterization procedure applied to Prelued in order to reach faster convergence proved to be ineffective (figure 5). This result should not be surprising given the simple mathematical formulation of the model, which does not allow the users to considerably change the parameter meaning and the model structure. Even if our approach allowed to reformulate four parameters out of six, this change in the parameters formulation did not lead to a substantial change in the overall model formulation, and this is likely to be the main reason of the ineffectiveness of this kind of procedure in this particular case.

It is not uncommon for data-based modeling exercises to show issues related to equifinality: frequently, the optimal parameter set is not uniquely defined. Instead, there may be many sets of parameters that all fit the data more or less equally well(Franks and Beven,

1997; Hollinger and Richardson, 2005; Schulz et al., 2001). This usually results in a delayed convergence, and can be due to high posterior correlation between parameters. These correlations could also be due to model overparameterization, which is known to lead to slow convergence (Rannala, 2002). An alternative solution to the issue of slow convergence was a two-steps calibration, using the posterior correlations between parameters resulting from the first step to reduce the number of parameters calibrated in the second step (thus reducing the dimensionality of the parameter space): this procedure allowed to reduce the number of parameters estimated, tackling both the issue of overparameterization and of equifinality, and reaching convergence with a smaller number of iterations (figure 7).

Since it was shown to be the same, the efficiency of the three considered algorithms should not drive their choice. The MHRW provided the parameter vector with best log-likelihood, but this did not result in better model performances over all. We suggest the DEMC algorithm as the best choice in this case study, due to the automatic computation of both the scale and orientation of the MCMC sampling. These are both user-defined in the MHRW algorithm, while only orientation is internally computed in the AM leaving scale as a user-defined setting. Since the optimal combination of scale and orientation is dependent on the prior distributions and on the data, the user might need several attempts to find it, making the calibration process even more time-consuming. It is also important to note that, once the optimal settings have been decided, the computational effort was the same for all the algorithms, even if other studies suggest that the DEMC algorithm is slower and requires more computational power than the others used in this case study(ter Braak, 2006).

5. Conclusions

In this study we compared the performance of three different Markov Chain Monte Carlo-based algorithms within a Bayesian framework to calibrate a Light Use Efficiency model (Prelued). The application of the three different algorithms of increasing complexity (Metropolis-Hastings Random Walk, Adaptive Metropolis, Differential Evolution Markov

Chain) with different number of iterations showed that all three MCMC-methods were similarly effective in reaching convergence. For all of them a very high number of iterations (10^6) is required for the Markov Chain to stabilize in the convergence region. This is due to the combination of at least two different factors: a strongly multiplicative mathematical structure, coupled with empirical parameters with neither ecological nor physiological meaning and parameters about which little is known in the literature.

An alternative solution to a very high number of iterations was a two-steps calibration, using the posterior correlations between parameters resulting from the first step to reduce the number of parameters calibrated in the second step. This approach reduced the computational effort necessary to reach proper convergence and was less time consuming than the previous one.

We suggest the DEMC algorithm as the best choice in this case study, even if its efficiency has proven to be similar to the other algorithms used, due to the automatic computation of both the scale and orientation of the MCMC sampling.

6. References

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Trans. Automat. Contr. 19, 716–723.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements. Irrig. Drain. Pap. 65 (300 pp.), United Nations Food Agric. Organ. Rome, Italy.

Andrieu, C., de Freitas, N., Doucet, A., Jordan, M., 2003. An Introduction to MCMC for Machine Learning. Mach. Learn. 50, 5–43. doi:10.1023/A:1020281327116

Asmussen, S., Glynn, P.W., 2007. Stochastic simulation: Algorithms and analysis. Springer Science & Business Media.

Baldocchi, D.D., 2008. "Breathing" of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. 56, 1–26. doi:10.1071/BT07151

Bernardo, J.M., Smith, A.F.M., 2009. Bayesian theory. John Wiley & Sons.

Blanco, J.A., Seely, B., Welham, C., Kimmins, J.P. (Hamish), Seebacher, T.M., 2007. Testing the performance of a forest ecosystem model (FORECAST) against 29 years of field data in a Pseudotsuga menziesii plantation. Can. J. For. Res. 37, 1808–1820. doi:10.1139/X07-041

Bloom, A.A., Williams, M., 2015. Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological" common sense" in a model–data fusion framework. Biogeosciences 12, 1299–1315.

Browne, W.J., Steele, F., Golalizadeh, M., Green, M.J., 2009. The use of simple reparameterizations to improve the efficiency of Markov chain Monte Carlo estimation for multilevel models with applications to discrete time survival models. J. R. Stat. Soc. Ser. A (Statistics Soc. 172, 579–598.

Bryson, A.E., 1975. Applied optimal control: optimization, estimation and control. CRC Press.

Buzzi-Ferraris, G., Manenti, F., 2010. Better reformulation of kinetic models. Comput. Chem. Eng. 34, 1904–1906.

Canadell, J.G., Le Quéré, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A., Marland, G., 2007. Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. Proc. Natl. Acad. Sci. 104 , 18866–18870. doi:10.1073/pnas.0702737104

Casella, G., Robert, C.P., 1999. Monte Carlo statistical methods.

Chevallier, F., Viovy, N., Reichstein, M., Ciais, P., 2006a. On the assignment of prior errors in Bayesian inversions of CO2 surface fluxes. Geophys. Res. Lett. 33, 1–5. doi:10.1029/2006GL026496

Chevallier, F., Viovy, N., Reichstein, M., Ciais, P., 2006b. On the assignment of prior errors in Bayesian inversions of CO2 surface fluxes. Geophys. Res. Lett. 33, 1–5. doi:10.1029/2006GL026496

Convertino, M., Muñoz-Carpena, R., Chu-Agor, M.L., Kiker, G.A., Linkov, I., 2014. Untangling drivers of species distributions: Global sensitivity and uncertainty analyses of MAXENT. Environ. Model. Softw. 51, 296–309. doi:10.1016/j.envsoft.2013.10.001

Davi, H., Dufrene, E., Francois, C., Lemaire, G., Loustau, D., Bosc, a, Rambal, S., Granier, a, Moors, E., 2006. Sensitivity of water and carbon fluxes to climate changes from 1960 to 2100 in European forest ecosystems. Agric. For. Meteorol. 141, 35–56.

Davidson, E.A., Richardson, A.D., Savage, K.E., Hollinger, D.Y., 2006. A distinct seasonal pattern of the ratio of soil respiration to total ecosystem respiration in a spruce-dominated forest. Glob. Chang. Biol. 12, 230–239. doi:0.1111/j.1365-2486.2005.01062.x

De Weirdt, M., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., Ciais, P., Steppe, K., 2012. Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model. Geosci. Model Dev. 5, 1091–1108. doi:10.5194/gmd-5-1091-2012

Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C., Wisniewski, J., 1994. Carbon pools and flux of global forest ecosystems. Science (80-.). 263, 185–190.

Dufrêne, E., Davi, H., François, C., Maire, G. Le, Dantec, V. Le, Granier, a., 2005. Modelling carbon and water cycles in a beech forestPart I: Model description and uncertainty analysis on modelled NEE. Ecol. Modell. 185, 407–436.

Duursma, R. a, Kolari, P., Perämäki, M., Pulkkinen, M., Mäkelä, A., Nikinmaa, E., Hari, P., Aurela, M., Berbigier, P., Bernhofer, C.H., Grünwald, T., Loustau, D., Mölder, 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 Physiol. 29, 621–639. doi:10.1093/treephys/tpp010

Ellison, A.M., 2004. Bayesian inference in ecology. Ecol. Lett. 7, 509–520.

Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecol. Appl. 6, 1036–1046.

Field, C.B., 2001. Plant Physiology of the "Missing" Carbon Sink. Plant Physiol. 125, 25–28.

Foley, J.A., Prentice, I.C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., Haxeltine, A., 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochem. Cycles 10, 603–628. doi:10.1029/96GB02692

Franks, S.W., Beven, K.J., 1997. Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. J. Geophys. Res. 102, 23,991–23,999.

Fu, Y.H., Campioli, M., Van Oijen, M., Deckmyn, G., Janssens, I.A., 2012. Bayesian comparison of six different temperature-based budburst models for four temperate tree species. Ecol. Modell. 230, 92–100.

Gao, C., Wang, H., Weng, E., Lakshmivarahan, S., Zhang, Y., Luo, Y., 2011. Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics. Ecol. Appl. 21, 1461–1473.

Ghazoul, J., McAllister, M., 2003. Communicating complexity and uncertainty in decision-making contexts: Bayesian approaches to forest research. Int. For. Rev. 5, 9–19.

Gilks, W., Roberts, G., 1996. Strategies for Improving MCMC., in: Gilks, W., Richardson, S., Spiegelhalter, D. (Eds.), Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, FL., pp. 89–114.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov chain Monte Carlo in practice. CRC press.

Goodale, C.L., Apps, M.J., Birdsey, R.A., Field, C.B., Heath, L.S., Houghton, R.A., Jenkins, J.C., Kohlmaier, G.H., Kurz, W., Liu, S., Nabuurs, G.-J., Nilsson, S., Shvidenko, A.Z., 2002. Forest Carbon Sinks In The Northern Hemisphere. Ecol. Appl. 12, 891–899. doi:10.1890/1051-0761(2002)012[0891:FCSITN]2.0.CO;2

Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote Sens. Environ. 70, 29–51. doi:10.1016/S0034-4257(99)00056-5

Grant, R.F., Baldocchi, D.D., Ma, S., 2012. Ecological controls on net ecosystem productivity of a seasonally dry annual grassland under current and future climates: Modelling with ecosys. Agric. For. Meteorol. 152, 189–200.

Grassi, G., Magnani, F., 2005. Stomatal, mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees. Plant, Cell Environ. 28, 834–849. doi:10.1111/j.1365-3040.2005.01333.x

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli 7, 223–242.

Häkkinen, R., Linkosalo, T., Hari, P., 1998. Effects of dormancy and environmental factors on timing of bud burst in Betula pendula. Tree Physiol. 18, 707–712. doi:10.1093/treephys/18.10.707

Hill, T.C., Ryan, E., Williams, M., 2012. The use of CO_2 flux time series for parameter and carbon stock estimation in carbon cycle research. Glob. Chang. Biol. 18, 179–193.

Hollinger, D.Y., Richardson, A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol. 25, 873–885. doi:10.1093/treephys/25.7.873

Horn, J.E., Schulz, K., 2011. Identification of a general light use efficiency model for gross primary production. Biogeosciences 8, 999–1021. doi:10.5194/bg-8-999-2011

Horn, J.E., Schulz, K., 2011. Spatial extrapolation of light use efficiency model parameters to predict gross primary production. J. Adv. Model. Earth Syst. 3, 21:

Huntington, T.G., 2006. Evidence for intensification of the global water cycle: Review and synthesis. J. Hydrol. 319, 83–95. doi:http://dx.doi.org/10.1016/j.jhydrol.2005.07.003

Jansson, P.-E., Svensson, M., Kleja, D.B., Gustafsson, D., 2008. Simulated climate change impacts on fluxes of carbon in Norway spruce ecosystems along a climatic transect in Sweden. Biogeochemistry 89, 81–94.

Jarvis, A.J., Stauch, V.J., Schulz, K., Young, P.C., 2004. The seasonal temperature dependency of photosynthesis and respiration in two deciduous forests. Glob. Chang. Biol. 10, 939–950. doi:10.1111/j.1529-8817.2003.00743.x.

Jaynes, E.T., 2003. Probability theory: the logic of science. Cambridge university press.

Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., Saugier, B., 2008. Nitrogen controls plant canopy light-use efficiency in temperate and boreal ecosystems. J. Geophys. Res. 113, 1–19. doi:10.1029/2007JG000676

Kimmins, J.P., 1986. Predicting the consequences of intensive forest harvesting on long-term productivity: the need for a hybrid model such as FORCYTE-11. Rapp. Institutionen Ekol. och Miljovard, Sveriges Lantbruksuniversitet 31–84.

Kimmins, J.P., Blanco, J.A., Seely, B., Welham, C., Scoullar, K., 2008. Complexity in modelling forest ecosystems: How much is enough? For. Ecol. Manage. 256, 1646–1658. doi:10.1016/j.foreco.2008.03.011

Kimmins, J.P., Mailly, D., Seely, B., 1999. Modelling forest ecosystem net primary production: the hybrid simulation approach used in forecast. Ecol. Modell. 122, 195–224. doi:10.1016/S0304-3800(99)00138-6

Kramer, K., Leinonen, I., Bartelink, H.H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, a. J., Froer, O., Gracia, C. a., Granier, a., Grunwald, T., Hari, P., Jans, W., Kellomaki, S., Loustau, D., Magnani, F., Markkanen, T., Matteucci, G., Mohren, G.M.J., Moors, E., Nissinen, a., Peltola, H., Sabate, S., Sanchez, a., Sontag, M., Valentini, R., Vesala, T., 2002. Evaluation of six process-based forest growth models using eddy-covariance measurements of CO_2 and H_2O fluxes at six forest sites in Europe. Glob. Chang. Biol. 8, 213–230. doi:10.1046/j.1365-2486.2002.00471.x

Kumar, M., Monteith, J.L., 1981. Remote sensing of crop growth, in: Plants and the Daylight Spectrum. Academic Press: London, pp. 133–144.

Landsberg, J., 2003. Modelling forest ecosystems: state of the art, challenges, and future directions. Can. J. For. Res. 33, 385–397. doi:10.1139/x02-129

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

Landsberg, J., Waring, R.H., Coops, N.C., 2003. Performance of the forest productivity model 3-PG applied to a wide range of forest types. For. Ecol. Manage. 172, 199–214. doi:10.1016/S0378-1127(01)00804-0

Magnani, F., Grace, J., Borghetti, M., 2002. Adjustment of tree structure in response to the environment under hydraulic constraints. Funct. Ecol. 16, 385–393. doi:10.1046/j.1365-2435.2002.00630.x

Mäkelä, A., 1988. Performance analysis of a process-based stand growth model using Monte Carlo techniques. Scand. J. For. Res. 3, 315–331.

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, 369–76. doi:10.1093/treephys/24.4.369

Mäkelä, A., Kolari, P., Karimäki, J., Nikinmaa, E., Perämäki, M., Hari, P., 2006. Modelling five years of weather-driven variation of GPP in a boreal forest. Agric. For. Meteorol. 139, 382–398. doi:0.1016/j.agrformet.2006.08.017

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen, P., 2000a. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20, 289–298.

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen, P., 2000b. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20, 289–298.

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. Chang. Biol. 14, 92–108. doi:10.1111/j.1365-2486.2007.01463.x

Marcolla, B., Cescatti, A., Montagnani, L., Manca, G., Kerschbaumer, G., Minerbi, S., 2005. Importance of advection in the atmospheric CO2 exchanges of an alpine forest. Agric. For. Meteorol. 130, 193–206. doi:10.1016/j.agrformet.2005.03.006

McCallum, I., Franklin, O., Moltchanova, E., Merbold, L., Schmullius, C., Shvidenko, A., Schepaschenko, D., Fritz, S., 2013. Improved light and temperature responses for light-use-efficiency-based GPP models. Biogeosciences 10, 6577–6590. doi:10.5194/bg-10-6577-2013

McMurtrie, R.E., Gholz, H.L., Linder, S., Gower, S.T., 1994. Climatic factors controlling the productivity of pine stands : a model-based analysis. Ecol. Bull. 43, 173–188.

Medlyn, B., Barrett, D., Landsberg, J., Sands, P., Clement, R., 2003. Conversion of canopy intercepted radiation to photosynthate: review of modelling approaches for regional scales. Funct. Plant Biol. 30, 153–169. doi:10.1071/FP02088

Melaas, E.K., Richardson, A.D., Friedl, M.A., Dragoni, D., Gough, C.M., Herbst, M., Montagnani, L., Moors, E., 2013. Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. Agric. For. Meteorol. 171-172, 46–56. doi:110.1016/j.agrformet.2012.11.018

Milne, A.E., Glendining, M.J., Bellamy, P., Misselbrook, T., Gilhespy, S., Casado, M.R., Hulin, A., Van Oijen, M., Whitmore, A.P., 2014. Analysis of uncertainties in the estimates of nitrous oxide and methane emissions in the UK's greenhouse gas inventory for agriculture. Atmos. Environ. 82, 94–105.

Minunno, F., Xenakis, G., Perks, M.P., Mencuccini, M., 2010. Calibration and validation of a simplified process- based model for the prediction of the carbon balance of Scottish Sitka spruce (Picea sitchensis) plantations. Can. J. For. Res. 40, 2411–2426. doi:10.1139/X10-181

Mo, X., Chen, J.M., Ju, W., Black, T.A., 2008. Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an ensemble Kalman filter. Ecol. Modell. 217, 157–173. doi:10.1016/j.ecolmodel.2008.06.021

Montagnani, L., Manca, G., Canepa, E., Georgieva, E., Acosta, M., Feigenwinter, C., Janous, D., Kerschbaumer, G., Lindroth, A., Minach, L., Minerbi, S., Mölder, M., Pavelka, M., Seufert, G., Zeri, M., Ziegler, W., 2009. A new mass conservation approach to the study

of CO2 advection in an alpine forest. J. Geophys. Res. Atmos. 114, D07306. doi:10.1029/2008JD010650

Monteith, J.L., Moss, C.J., 1977. Climate and the Efficiency of Crop Production in Britain [and Discussion]. Philos. Trans. R. Soc. Lond. B 281, 277–294. doi:10.1098/rstb.1977.0140

Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., Liski, J., Zhou, L., Alexeyev, V., Hughes, M.K., 2001. A large carbon sink in the woody biomass of Northern forests. Proc. Natl. Acad. Sci. 98 , 14784–14789. doi:10.1073/pnas.261555198

Nightingale, J.M., Phinn, S.R., Held, A.A., 2004. Ecosystem process models at multiple scales for mapping tropical forest productivity. Prog. Phys. Geogr. 28, 241–281. doi:10.1191/0309133304pp411ra

Ogle, K., 2009. Hierarchical Bayesian statistics: merging experimental and modeling approaches in ecology. Ecol. Appl. 19, 577–581.

Ogle, K., Barber, J.J., 2008. Bayesian data—model integration in plant physiological and ecosystem ecology, in: Progress in Botany. Springer, pp. 281–311.

Oulehle, F., Evans, C.D., Hofmeister, J., Krejci, R., Tahovska, K., Persson, T., Cudlin, P., Hruska, J., 2011. Major changes in forest carbon and nitrogen cycling caused by declining sulphur deposition. Glob. Chang. Biol. n/a–n/a. doi:10.1111/j.1365-2486.2011.02468.x

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., 2012a. Does canopy mean nitrogen concentration explain variation in canopy light use efficiency across 14 contrasting forest sites? Tree Physiol. 32, 200–18. doi:10.1093/treephys/tpr140

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

Pretzsch, H., Grote, R., Reineking, B., Rötzer, T., Seifert, S., 2008. Models for forest ecosystem management: a European perspective. Ann. Bot. 101, 1065–1087. doi:10.1093/aob/mcm246

Rahn, K., Butterbach-bahl, K., Werner, C., 2011. Selection of likelihood parameters for complex models determines the effectiveness of Bayesian calibration. Ecol. Inform. 6, 333–340. doi:10.1016/j.ecoinf.2011.08.002

Rannala, B., 2002. Identifiability of Parameters in MCMC Bayesian Inference of Phylogeny. Syst. Biol. 51, 754–760. doi:10.1080/10635150290102429

Rannik, Ü., Kolari, P., Vesala, T., Hari, P., 2006. Uncertainties in measurement and modelling of net ecosystem exchange of a forest. Agric. For. Meteorol. 138, 244–257. doi:10.1016/j.agrformet.2006.05.007

Rastetter, E.B., Aber, J.D., Peters, D.P.C., Ojima, D.S., Burke, I.C., 2003. Using mechanistic models to scale ecological processes across space and time. Bioscience 53, 68–76.

Raupach, M.R., Rayner, P.J., Barrett, D.J., DeFries, R.S., Heimann, M., Ojima, D.S., Quegan, S., Schmullius, C.C., 2005. Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Glob. Chang. Biol. 11, 378–397. doi:10.1111/j.1365-2486.2005.00917.x

Reckhow, K.H., 1990. Bayesian inference in non-replicated ecological studies. Ecology 71, 2053–2059.

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. Chang. Biol. 11, 1424–1439. doi:10.1111/j.1365-2486.2005.001002.x

Richardson, A.D., Hollinger, D.Y., 2005. Statistical modeling of ecosystem respiration using eddy covariance data: Maximum likelihood parameter estimation, and Monte Carlo simulation of model and parameter uncertainty, applied to three simple models. Agric. For. Meteorol. 131, 191–208. doi:10.1016/j.agrformet.2005.05.008

Robinson, A.P., Froese, R.E., 2004. Model validation using equivalence tests. Ecol. Modell. 176, 349–358. doi:10.1016/j.ecolmodel.2004.01.013

Rodeghiero, M., Cescatti, A., 2005. Main determinants of forest soil respiration along an elevation/temperature gradient in the Italian Alps. Glob. Chang. Biol. 11, 1024–1041. doi:10.1111/j.1365-2486.2005.00963.x

Running, S.W., Baldocchi, D.D., Turner, D.P., Gower, S.T., Bakwin, P.S., Hibbard, K.A., 1999. A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sens. Environ. 70, 108–127.

Sacks, W.J., Schimel, D.S., Monson, R.K., Braswell, B.H., 2006. Model-data synthesis of diurnal and seasonal CO2 fluxes at Niwot Ridge, Colorado. Glob. Chang. Biol. 12, 240–259. doi:10.1111/j.1365-2486.2005.01059.x

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environ. Model. Softw. 25, 1508–1517. doi:10.1016/j.envsoft.2010.04.012

Saxton, K.E., Rawls, W.J., Romberger, J.S., Papendick, R.I., 1986. Estimating Generalized Soil-water Characteristics from Texture1. Soil Sci. Soc. Am. J. 50, 1031–1036. doi:10.2136/sssaj1986.03615995005000040039x

Schaefer, K., Schwalm, C.R., Williams, C., Arain, M.A., Barr, A., Chen, J.M., Davis, K.J., Dimitrov, D., Hilton, T.W., Hollinger, D.Y., 2012. A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. J. Geophys. Res. 117. doi:10.1029/2012JG001960

Schulz, K., Jarvis, A.J., Beven, K.J., Soegaard, H., 2001. The Predictive Uncertainty of Land Surface Fluxes in Response to Increasing Ambient Carbon Dioxide. J. Clim. 14, 2551–2562. doi:http://dx.doi.org/10.1175/1520-0442(2001)014<2551:TPUOLS>2.0.CO;2

Sims, D.A., Luo, H., Hastings, S., Oechel, W.C., Rahman, A.F., Gamon, J.A., 2006. Parallel adjustments in vegetation greenness and ecosystem CO2 exchange in response to drought in a Southern California chaparral ecosystem. Remote Sens. Environ. 103, 289–303. doi:10.1016/j.rse.2005.01.020

Sivia, D.S., 1996. Data Analysis: A Bayesian Tutorial. Oxford University Press, Oxford.

Smith, T.J., Marshall, L.A., 2008. Bayesian methods in hydrologic modeling : A study of recent advancements in Markov chain Monte Carlo techniques 44, 1–9. doi:10.1029/2007WR006705

Still, C.J., Randerson, J.T., Fung, I.Y., 2004. Large-scale plant light-use efficiency inferred from the seasonal cycle of atmospheric CO2. Glob. Chang. Biol. 10, 1240–1252. doi:10.1111/j.1365-2486.2004.00802.x

Storn, R., Price, K., 1997. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim. 11, 341–359. doi:10.1023/A:1008202821328

Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., Lindroth, a, 2008a. Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand. Ecol. Modell. 213, 331–344. doi:10.1016/j.ecolmodel.2008.01.001

Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., Lindroth, A., 2008b. Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand. Ecol. Modell. 213, 331–344. doi:10.1016/j.ecolmodel.2008.01.001

Ter Braak, C.J.F., 2006. A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. Stat. Comput. 16, 239–249. doi:10.1007/ s11222-006-8769-1

Theil, H., 1966. Applied Econometric Forecasting. North-Holland, Amsterdam.

Tjiputra, J., Roelandt, C., Bentsen, M., Lawrence, D., Lorentzen, T., Schwinger, J., Seland, O., Heinze, C., 2013. Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). Geosci. Model Dev. 6, 301–325. doi:10.5194/gmd-6-301-2013

Turner, D.P., Ritts, W.D., Styles, J.M., Yang, Z., Cohen, W.B., Law, B.E., Thornton, P.E., 2006. A diagnostic carbon flux model to monitor the effects of disturbance and interannual variation in climate on regional NEP. Tellus B 58, 476–490. doi:10.1111/j.1600-0889.2006.00221.x

Turner, D.P., Urbanski, S., Bremer, D., Wofsy, S.C., Meyers, T., Gower, S.T., Gregory, M., 2003. A cross-biome comparison of daily light use efficiency for gross primary production. Glob. Chang. Biol. 9, 383–395. doi:10.1046/j.1365-2486.2003.00573.x

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. doi:10.1016/j.agrformet.2011.06.017

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. doi:10.1016/j.agrformet.2011.06.017

Van Oijen, M., Cannell, M.G.R., Levy, P.E., 2004. Modelling biogeochemical cycles in forests: state of the art and perspectives, in: Andersson, F., Birot, Y., Päivinen, R. (Eds.), Towards the Sustainable Use of Europe's Forests—forest Ecosystem and Landscape Research: Scientific Challenges and Opportunities. European For. Inst., Joensuu, Finland, pp. 157–169.

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, 915–27. doi:10.1093/treephys/25.7.915

Vanclay, J.K., Skovsgaard, J.P., 1997. Evaluating forest growth models. Ecol. Modell. 98, 1–12. doi:10.1016/S0304-3800(96)01932-1

Veroustraete, F., Patyn, J., Myneni, R.B., 1994. Forcing of a simple ecosystem model with fAPAR and climatic data to estimate regional scale photosynthetic assimilation. Veg. Model. Clim. Chang. Eff. 151–177.

Vörösmarty, C.J., Sahagian, D., 2000. Anthropogenic disturbance of the terrestrial water cycle. Bioscience 50, 753–765.

Waring, R.H., Running, S.W., 1998. Forest Ecosystems: Analysis at Multiple Scales. Academic Press, San Diego.

Wellek, S., 2003. Testing Statistical Hypotheses of Equivalence. Chapman & Hall, London.

White, J.D., Running, S.W., 1994. Testing scale dependent assumptions in regional ecosystem simulations. J. Veg. Sci. 5, 687–702. doi:10.2307/3235883

White, M. a., Thornton, P.E., Running, S.W., Nemani, R.R., 2000. Parameterization and Sensitivity Analysis of the BIOME–BGC Terrestrial Ecosystem Model: Net Primary Production Controls. Earth Interact. 4, 1–85.

Wikle, C.K., Berliner, L.M., 2007. A Bayesian tutorial for data assimilation. Phys. D Nonlinear Phenom. 230, 1–16.

Williams, M., Law, B.E., Anthoni, P.M., H, M., 2001. Use of a simulation model and ecosystem flux data to examine carbon − water interactions in ponderosa pine. Tree Physiol. 287–298.

Wisskirchen, K., Tum, M., Gunther, K., Niklaus, M., Eisfelder, C., Knorr, W., 2013. Quantifying the carbon uptake by vegetation for Europe on a 1 km(2) resolution using a remote sensing driven vegetation model. Geosci. Model Dev. 6, 1623–1640.

Wu, Y., Liu, S., 2012. Automating calibration, sensitivity and uncertainty analysis of complex models using the R package Flexible Modeling Environment (FME): SWAT as an example. Environ. Model. Softw. 31, 99–109. doi:10.1016/j.envsoft.2011.11.013

Xenakis, G., Ray, D., Mencuccini, M., 2008. Sensitivity and uncertainty analysis from a coupled 3-PG and soil organic matter decomposition model. Ecol. Modell. 219, 1–16. doi:10.1016/j.ecolmodel.2008.07.020

Xiao, X., Zhang, Q., Hollinger, D., 2005. Modeling gross primary production of an evergreen needleleaf forest using MODIS and climate data. Ecol. Appl. 15, 954–969. doi:10.1890/04-0470

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L.L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A.H., Goulden, M.L., Hollinger, D.Y., Hu, Y., Law, B.E., Stoy, P.C., Vesala, T., Wofsy, S.C., 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agric. For. Meteorol. 143, 189–207. doi:10.1016/j.agrformet.2006.12.001

Zhang, X., Kondragunta, S., 2006. Estimating forest biomass in the USA using generalized allometric models and MODIS land products. Geophys. Res. Lett. 33, 1–5.

Zhu, G.F., Li, X., Su, Y.H., Zhang, K., Bai, Y., Ma, J.Z., Li, C.B., Hu, X.L., He, J.H., 2014. Simultaneously assimilating multivariate data sets into the two-source evapotranspiration model by Bayesian approach: application to spring maize in an arid region of northwestern China. Geosci. Model Dev. 7, 1467–1482. doi:10.5194/gmd-7-1467-2014

7. Tables and figures

Table 1.Prior probability distributions for each parameter in the Prelued model.

Parameter	Unit	Prior min.	Prior max.
β	gC mol^{-1}	0.0	1.5
γ	m^2 mol^{-1}	0.0	0.1
κ	kPa^{-1}	-10.0	0.0
X_0	°C	-100.0	0.0
τ	°C	0.0	100.0
S_{max}	°C	0.0	100.0

Table 2. Best parameter sets and log-likelihood values for the three MCMC algorithms applied (single-step calibration of Prelued), compared with the optimized parameter values found by Mäkelä et al. (2008) in similar ecosystems.

Site	Year	Dominant species	Algorithm	Best parameter vector / Optimized parameter value						Log-likelihood	Reference
				β	γ	κ	X_0	τ	S_{max}		
Lavarone	2004	*A. alba + P. abies*	MHRW	0.55	0.02	-0.92	-7.01	9.51	13.28	-117.78	-
			AM	0.56	0.02	-0.93	-6.89	9.19	12.91	-124.41	
			DEMC	0.56	0.02	-0.93	-6.60	9.52	12.21	-134.14	
Norunda	1999	*P. abies+ P. sylvestris*	-	0.49	0.002	-0.39	-10.0	5.0	29.0	-	Mäkelä et al. (2008)
Tharandt	2003	*P. abies*		0.66	0.016	-0.70	-5.0	2.0	19.50		
Bray	2001	*P. pinaster*		0.49	0.021	-0.06	-1.0	2.0	19.0		

Table 3. Posterior coefficients of correlation between parameters (single-step calibration of Prelued). Coefficients that differ more than one order of magnitude or have different sign between algorithms are highlighted in bold text.

Algorithm	Parameter	β	γ	κ	X_0	τ	S_{max}
MHRW		1	0.91	0.135	-0.15	-0.262	0.369
AM	β	1	0.89	**0.039**	-0.095	-0.269	0.294
DEMC		1	0.896	0.156	-0.106	-0.257	0.291
MHRW		0.91	1	0.471	-0.13	-0.226	0.325
AM	γ	0.89	1	0.417	-0.106	-0.218	0.272
DEMC		0.896	1	0.512	-0.116	-0.263	0.27
MHRW		0.135	0.471	1	**0.006**	0.012	0.072
AM	κ	**0.039**	0.417	1	-0.021	0.067	0.062
DEMC		0.156	0.512	1	-0.017	**-0.07**	0.081
MHRW		-0.15	-0.13	**0.006**	1	0.434	-0.923
AM	X_0	-0.095	-0.106	-0.021	1	0.483	-0.928
DEMC		-0.106	-0.116	-0.017	1	0.418	-0.926
MHRW		-0.262	-0.226	0.012	0.434	1	-0.512
AM	τ	-0.269	-0.218	0.067	0.483	1	-0.578
DEMC		-0.257	-0.263	**-0.07**	0.418	1	-0.529
MHRW		0.369	0.325	0.072	-0.923	-0.512	1
AM	S_{max}	0.294	0.272	0.062	-0.928	-0.578	1
DEMC		0.291	0.27	0.081	-0.926	-0.529	1

Table 4. Means and variances of the 1st and 2nd half of last 50% of the longest chain for each algorithm (single-step calibration of Prelued).

Algorithm	Parameter	Mean of 1st half	Mean of 2nd half	Variance of 1st half	Variance of 2nd half
MHRW	β	0.55036	0.54845	0.0011	0.00116
	γ	0.01799	0.01779	0.00001	0.00001
	κ	0.924	0.92421	0.00816	0.00816
	X_0	6.96082	7.0221	0.54956	0.59826
	τ	9.60252	9.60196	1.96344	2.06421
	S_{max}	13.15563	13.24499	1.74995	1.92255
AM	β	0.56413	0.56479	0.00105	0.00101
	γ	0.01903	0.01898	0.00001	0.00001
	κ	0.92784	0.93448	0.00752	0.00758
	X_0	6.89452	6.90664	0.53949	0.53714
	τ	9.18685	9.07216	1.97048	1.82252
	S_{max}	12.93946	12.9531	1.55456	1.57706
DEMC	β	0.55799	0.55883	0.00083	0.00076
	γ	0.01849	0.0187	0.00001	0.00001
	κ	0.92787	0.92596	0.00675	0.00689
	X_0	6.7482	6.75782	0.45696	0.35082
	τ	9.54574	9.26533	1.35935	1.26707
	S_{max}	12.46423	12.52407	1.24835	1.08612
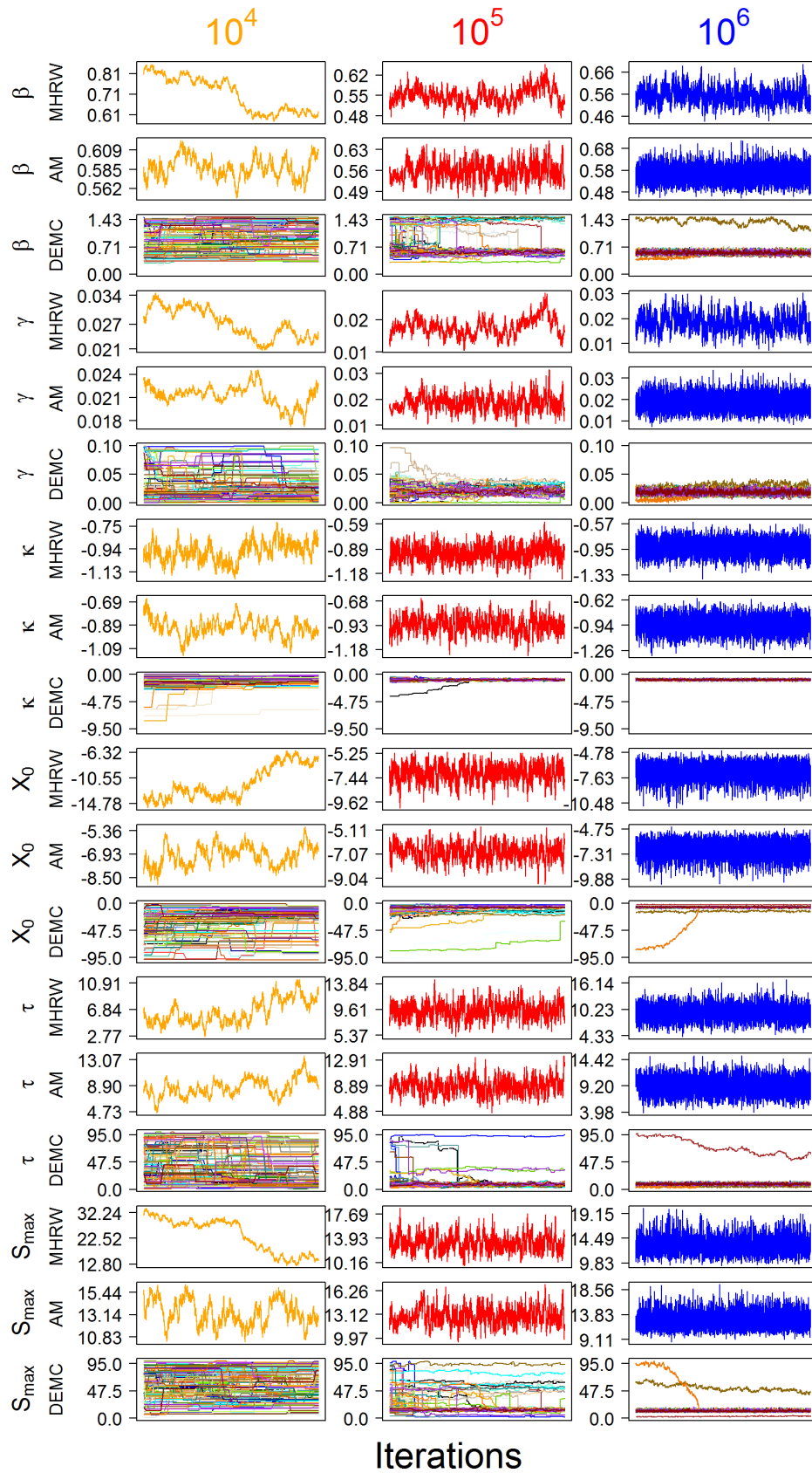
Figure 1.Traceplots of the post burn-in MCMC sampling for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the single-step calibration of the Prelued model.
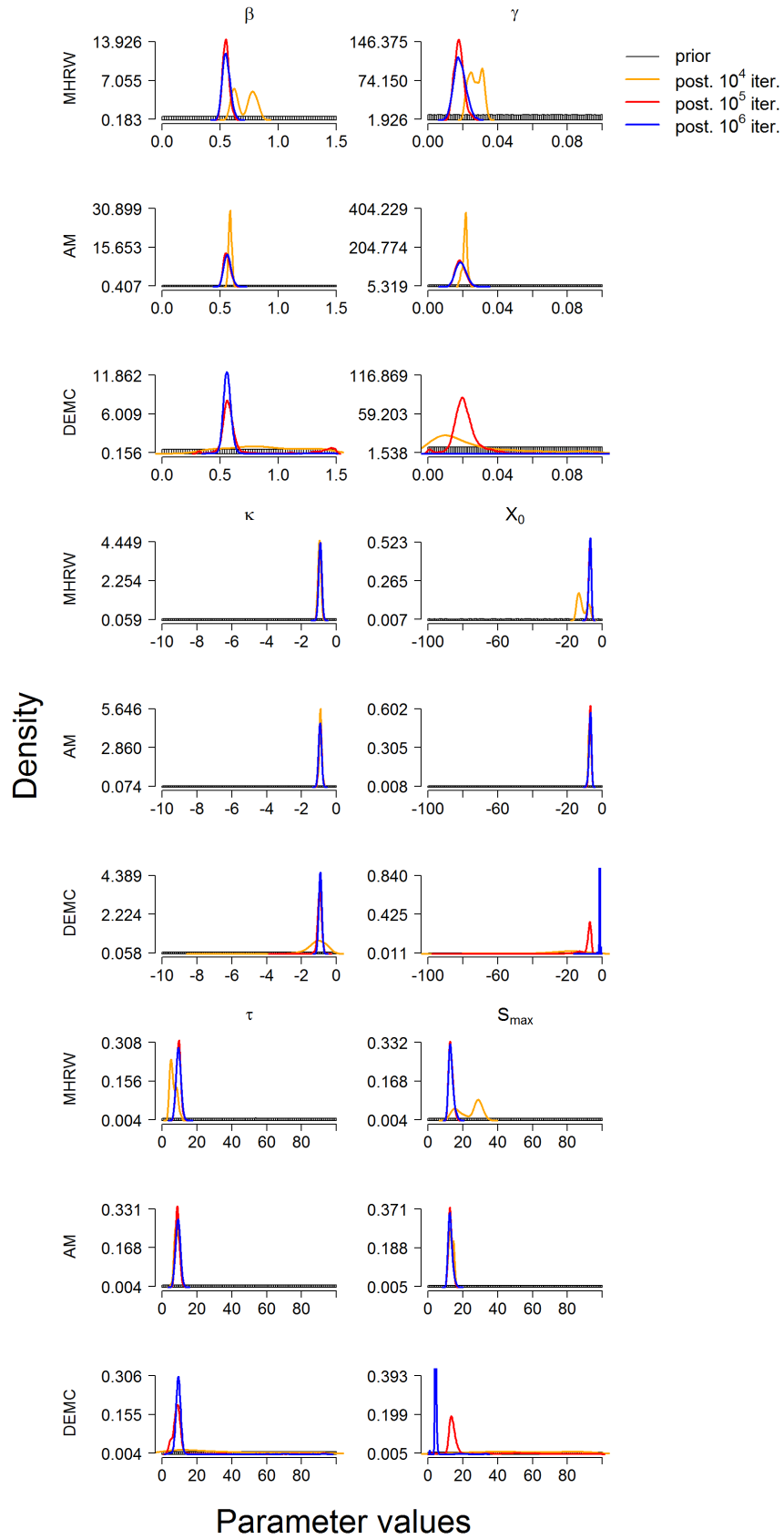
Figure 2.Posterior probability distributions of parameters for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the single-step calibration of the Prelued model.
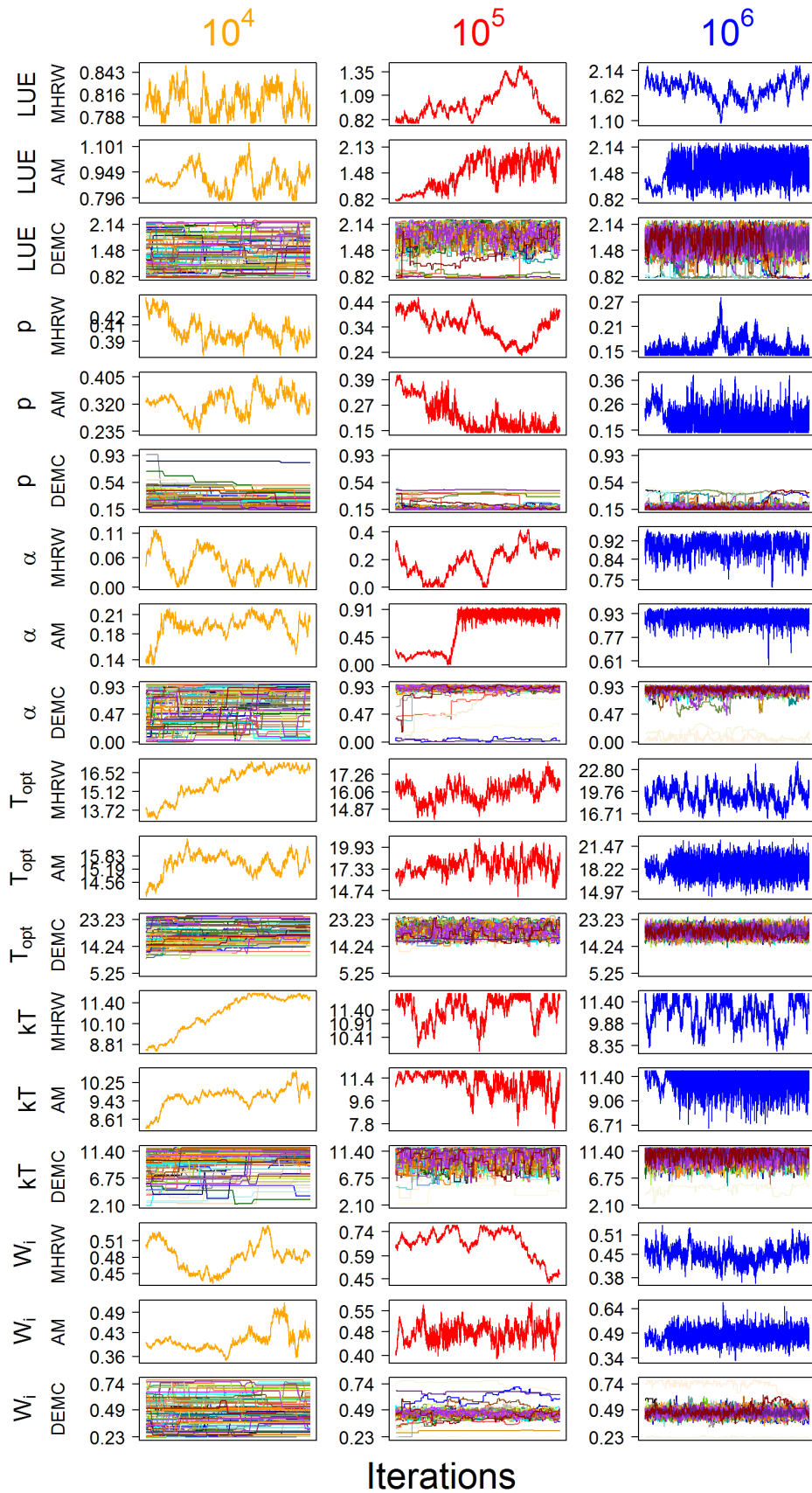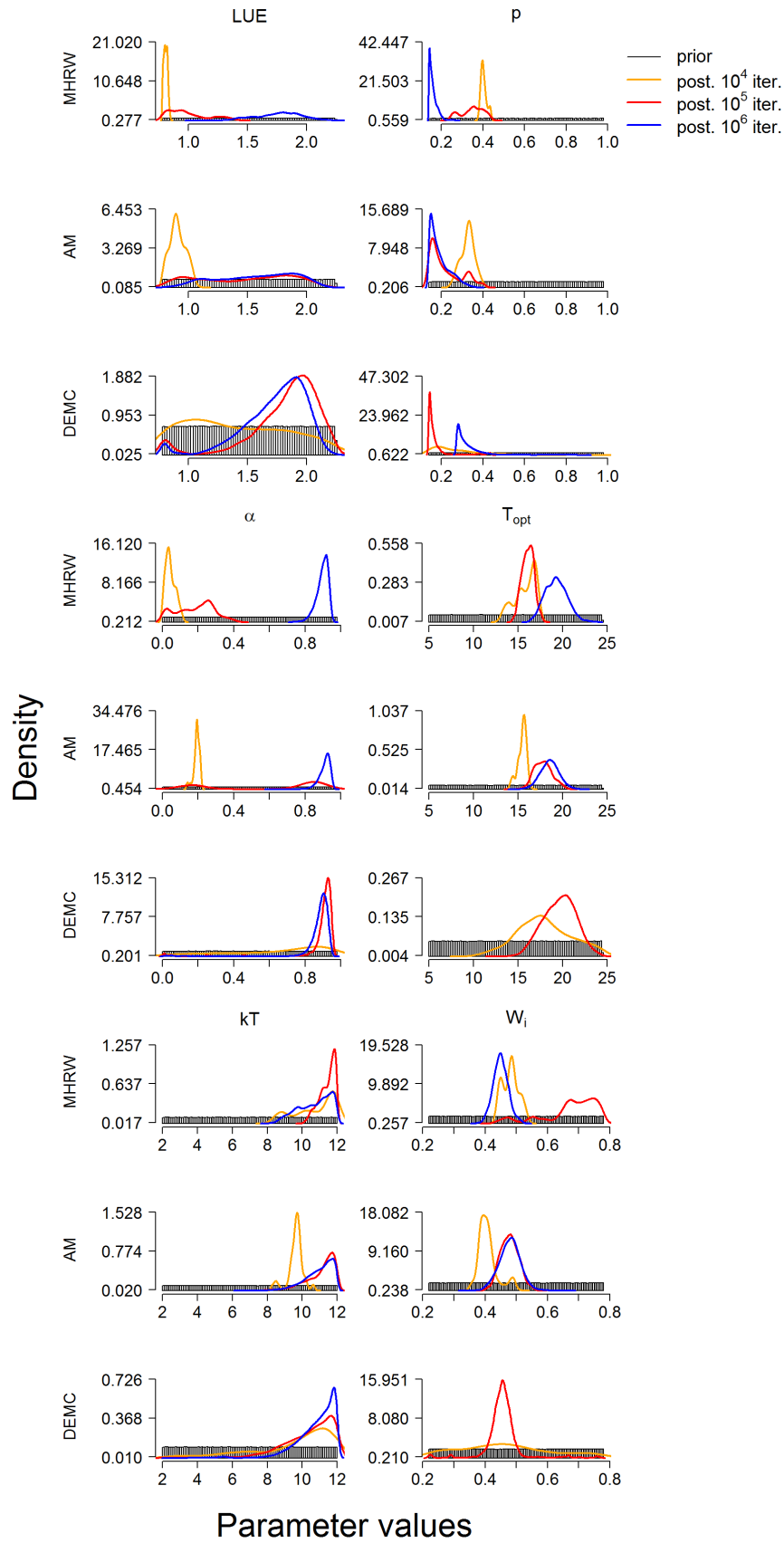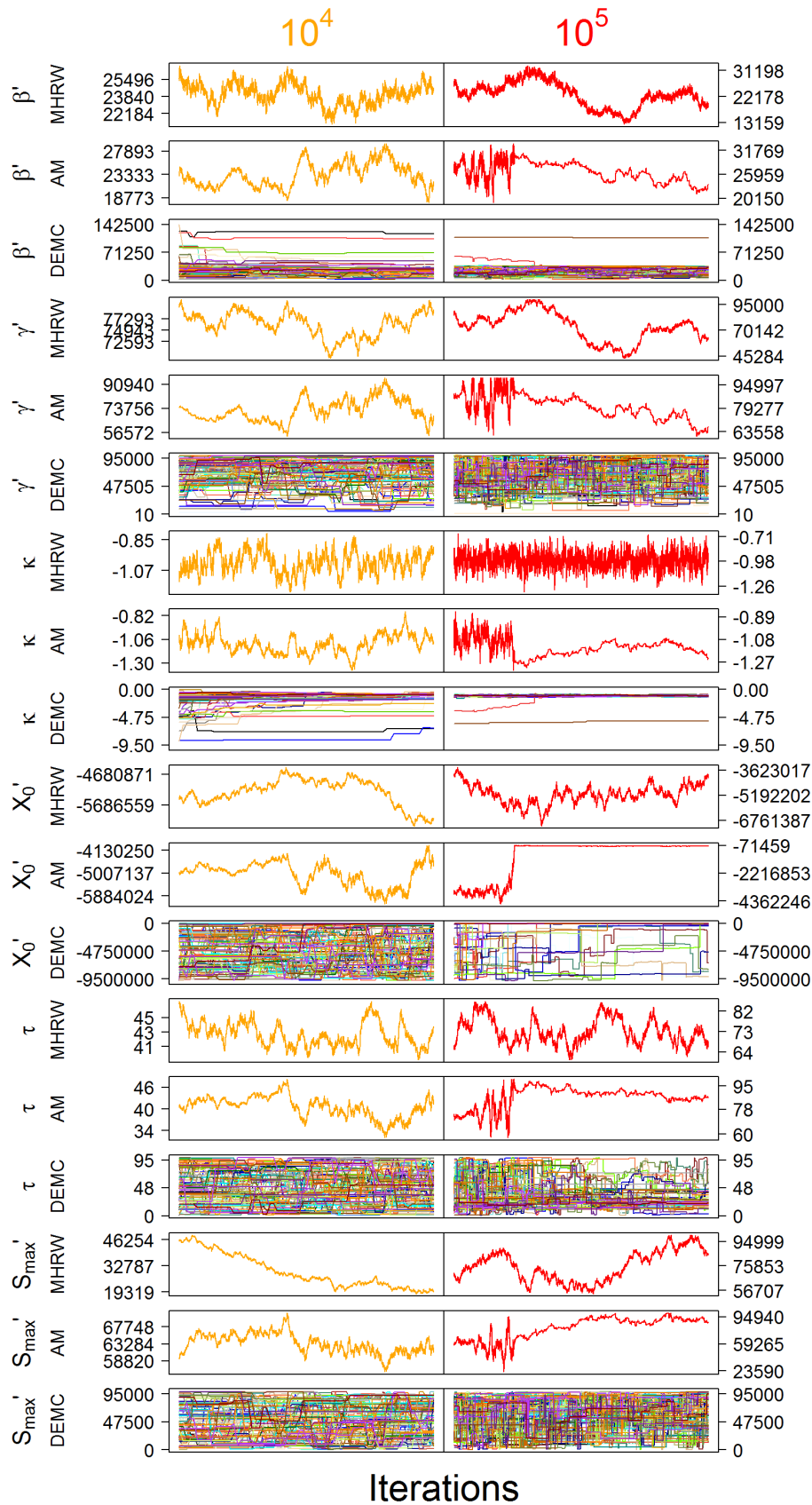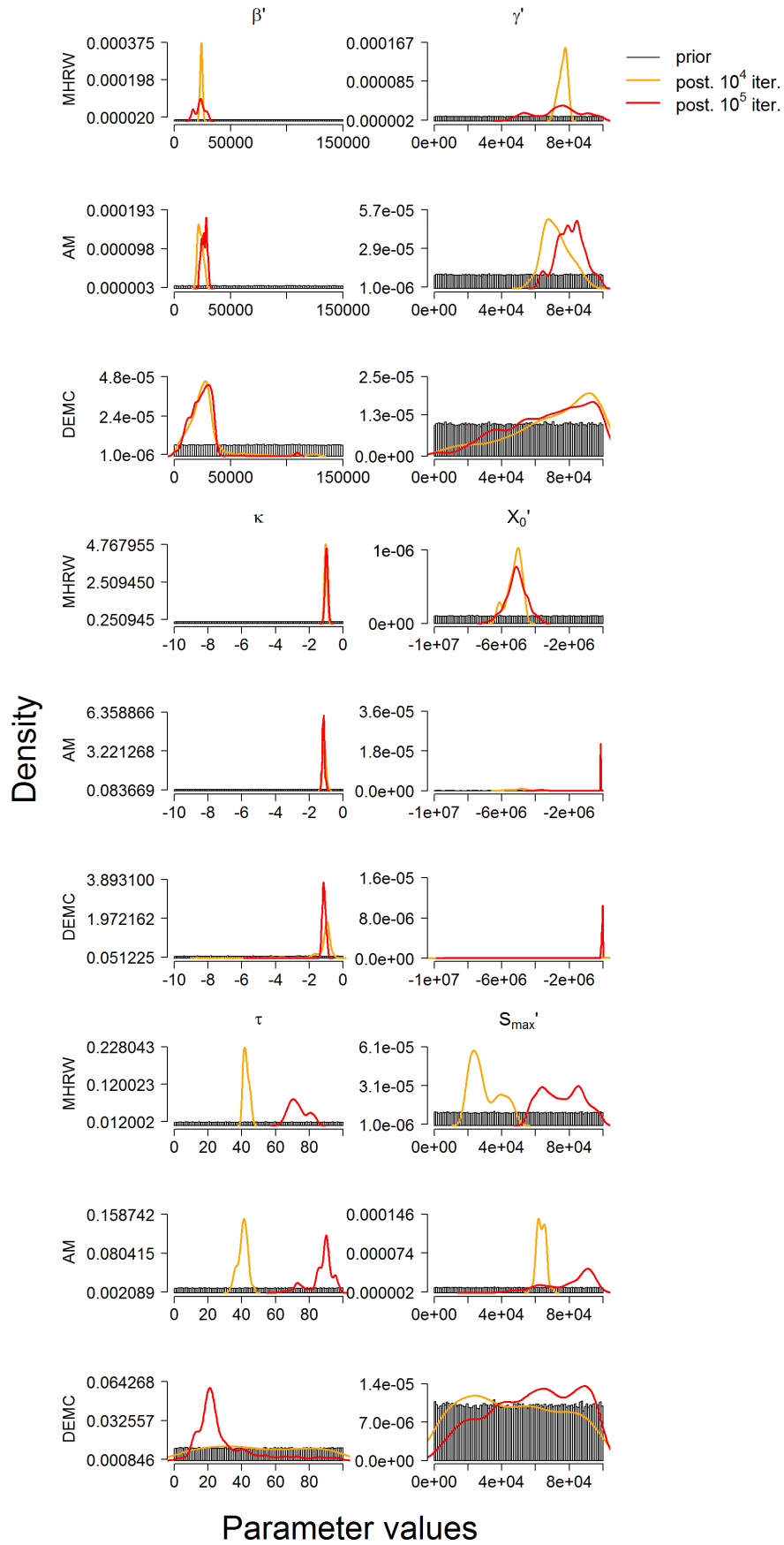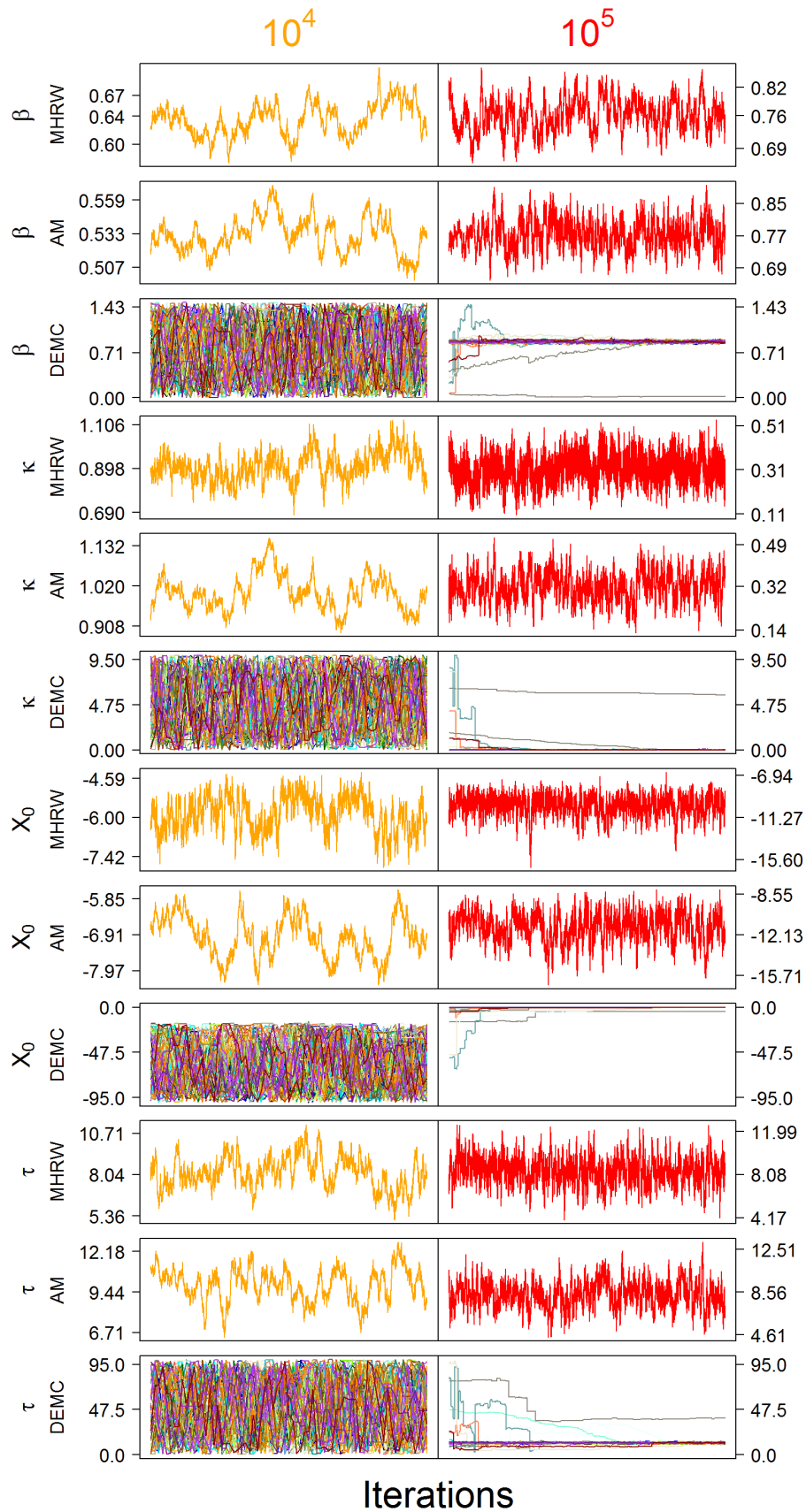
Figure 3. Traceplots of the post burn-in MCMC sampling for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the model by Horn and Schulz(2011b).

Figure 4. Posterior probability distributions of parameters for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the model by Horn and Schulz(2011b).

Figure 5. Traceplots of the post burn-in MCMC sampling for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the reparameterized Prelued model.
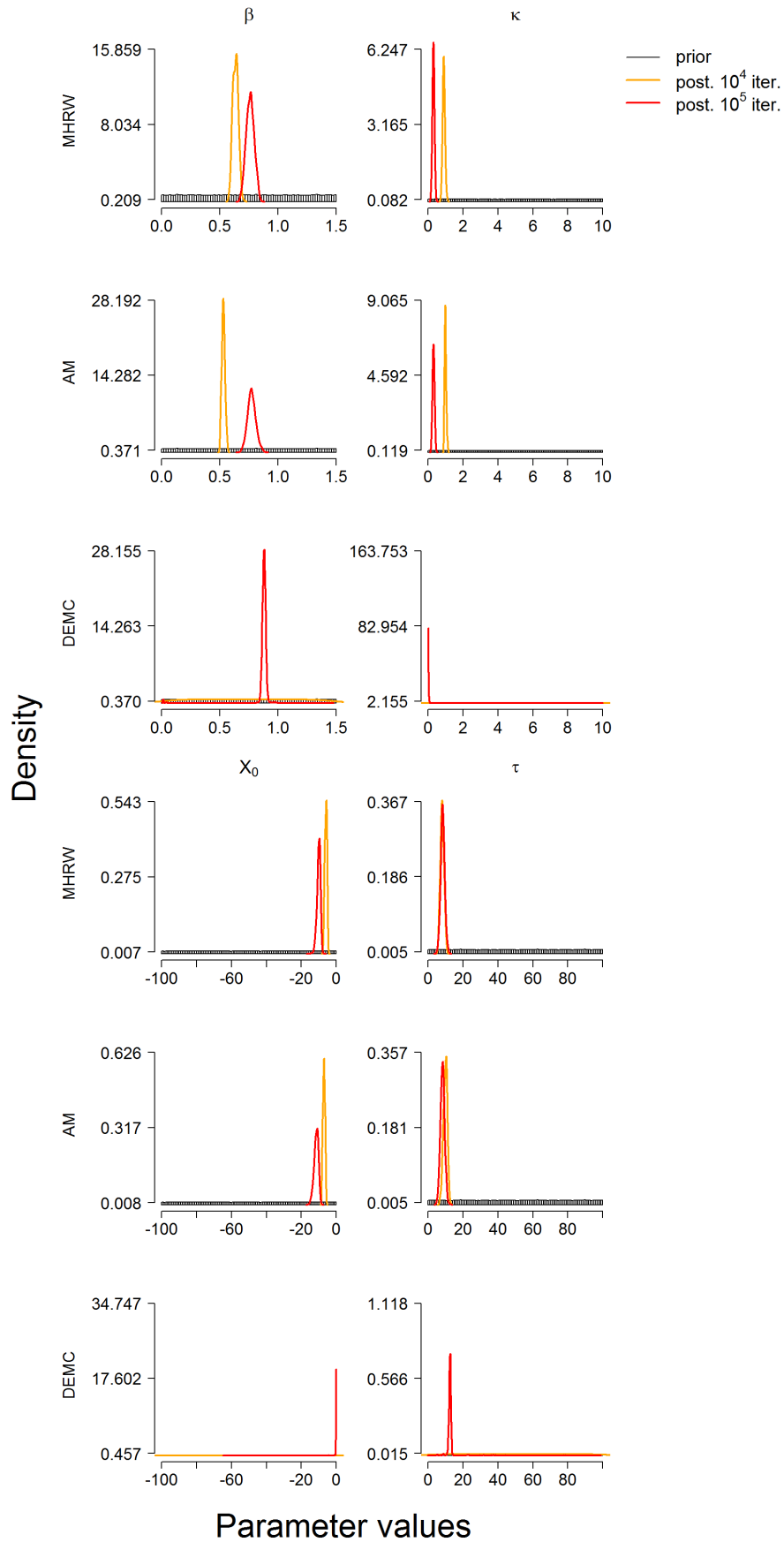
Figure 6.Posterior probability distributions of parameters for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the reparameterized Prelued model.

Figure 7. Traceplots of the post burn-in MCMC sampling for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the two-steps calibration of the Prelued model.
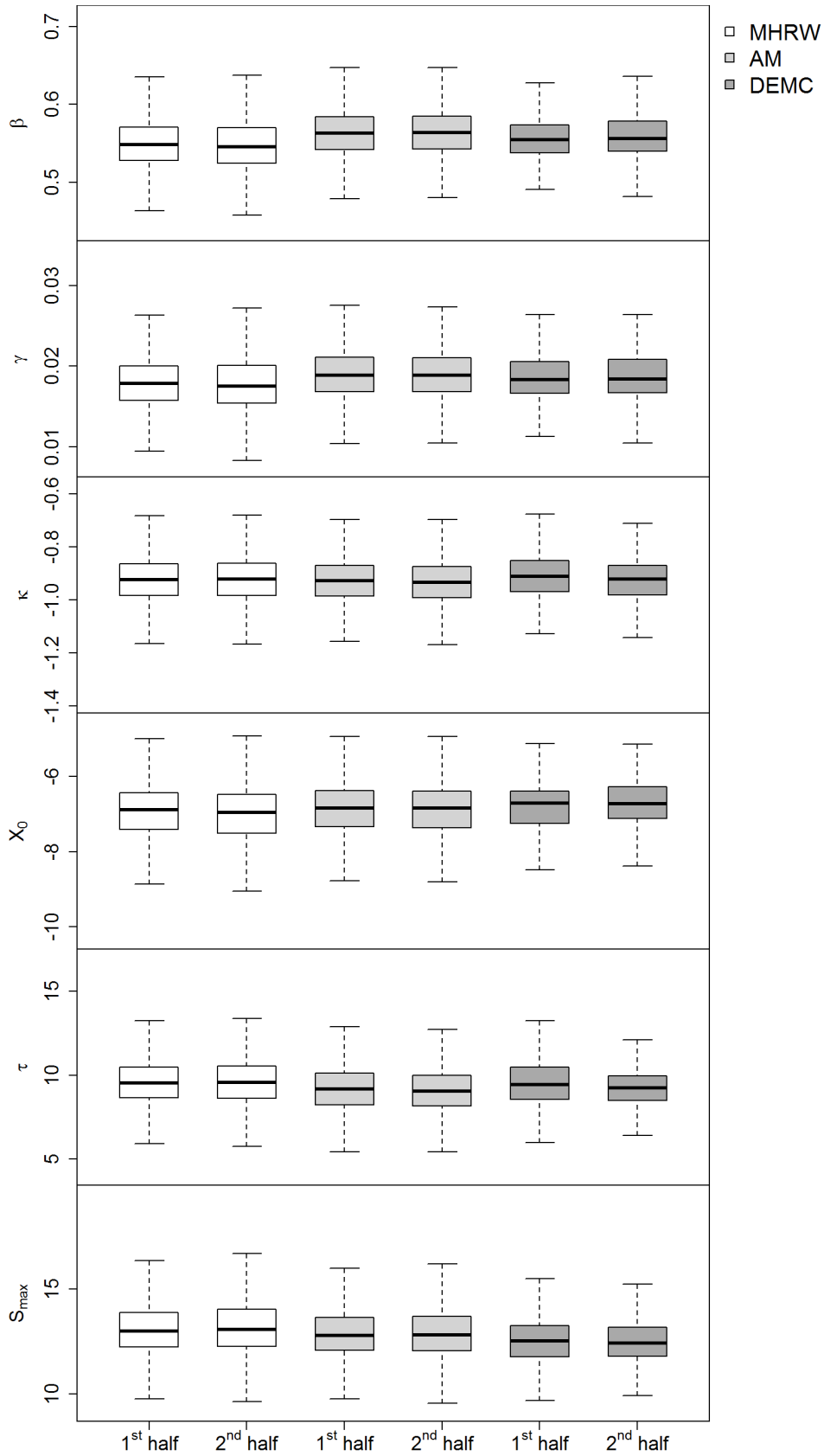
Figure 8.Posterior probability distributions of parameters for all the applied algorithms(MHRW, AM, DEMC) with different number of iterations, for the two-steps calibration of the Prelued model.

Figure 9.Boxplot of the 1st and 2nd half of the last 50% of the longest chain for each algorithm for the single-step calibration of the Prelued model.

Chapter 4: Bayesian calibration of the process-based forest model HYDRALL to a Silver Fir ecosystem in Northern Italy

Abstract

Forest models are being increasingly used to study ecosystem functioning, through the reproduction of carbon fluxes and productivity in very different forests all over the world. In this study we applied the physiology-based forest model HYDRALL to an alpine coniferous forest dominated by Silver fir in the Western Italian Alps, with a Bayesian approach to calibration based on a multiple-chains Markov Chain Monte Carlo sampling. The model was calibrated twice: once for daily Net Ecosystem Productivity (NEP), and once for both daily NEP and daily evapotranspiration (ET), to estimate the importance of incorporating additional data streams in the calibration procedure and to study the effects on its effectiveness. The model uncertainties on both NEP and ET were estimated, as well as the uncertainties on calibrated model parameters, for each calibration procedure.

The accuracy of the model estimates of daily NEP and ET, as well as the model uncertainties and the calibrated parameter values and their uncertainties differed widely depending on the variables used during the calibration process, with good model performance on NEP but poor on ET when calibrate only on NEP, and good model performance on both NEP and ET when ET was added as a calibrating variable. Model uncertainties were always smaller than data uncertainties, with variable magnitude depending on the calibration procedure applied. The calibrated parameter values and their distributions also varied, with 7 parameters out of 10 showing a narrowed peaked posterior distribution when calibrating only on NEP, and all parameters showing a narrowed peaked posterior distribution when calibrating both on NEP and ET. In conclusion, the HYDRALL model appears to be suitable for estimating daily NEP and ET in this type of ecosystem, when calibrated on more than one variable, and ET appears to be a key model output to the success of the calibration procedure. The Bayesian calibration resulted in good model performances in respect to other methods used in previous studies, and allowed us to estimate uncertainties on both parameter values

and model estimates, which were useful to analyse more in detail the impact of additional variables in the calibration procedure.

1. Introduction

Biological, chemical and physical processes controlling trace gas fluxes and their response to environmental perturbations are the subject of extensive environmental research. Central to this research is the design, construction and application of mathematical models that compute rates of biosphere–atmosphere trace gas exchange (Baldocchi and Meyers, 1998). To understand the mechanisms of the continual rise in atmospheric CO_2 and its consequences, it is important to investigate the global carbon cycle, and its interactions with the fluxes of energy and mass between the biosphere, lithosphere, the oceans and the atmosphere. The evaluation and prediction of changes in the carbon dynamics at the ecosystem level has been a key issue in studies of global climate change (Friend et al., 1993; Landsberg and Gower, 1997). In the global Carbon (C) budget, the terrestrial biosphere is thought to be a significant C sink of the order of 2.8 PgC yr^{-1} (Canadell et al., 2007; Field, 2001; Myneni et al., 2001). Although the biomes responsible of this C sequestration are still uncertain, it appears likely that temperate and boreal forest ecosystems of the northern hemisphere give a considerable contribution to this process, with approximately 0.6-0.7 PgC yr^{-1} (Goodale et al., 2002). Therefore, the understanding of forest ecosystem processes, their magnitude, their variation in time and space, are amongst the main objectives of ecological studies.

Over the past two decades, interest in issues such as climate change and global carbon cycles has increased development and application of ecosystem scale primary production models (White et al., 2000). During this period, a number of models based on biochemical and physiological processes have been developed. One of the critical aspects in the application of any forest model is the estimation of the model parameters and of the uncertainty associated with them (van Oijen et al., 2005). For this purpose we considered the application of the Bayesian approach to calibration, a method that has become more and more popular in the last few years to obtain insights on both

model predictions and uncertainties (Richardson et al., 2010; van Oijen and Thomson, 2010; van Oijen et al., 2005). The main characteristic of a Bayesian calibration is that it quantifies model inputs and outputs in the form of probability distributions, and applies the rules of probability theory to update the distributions when new data are obtained (Sivia, 1996; van Oijen et al., 2005). Bayesian approaches have been much less frequently used in forest research (Ghazoul and McAllister, 2003) than in other environmental sciences (Ellison, 2004, 1996; Reckhow, 1990), likely because the Bayesian approach of analyzing everything probabilistically, even if the model has many different inputs and outputs, makes this calibration computationally demanding. This has hampered its practical use in the past, and a very few examples exist for forest models with many parameters (Svensson et al., 2008). However, recent developments in sampling-based evaluation of probability distributions, in particular Markov Chain Monte Carlo techniques, can help alleviate the computational problem (van Oijen et al., 2005).

Models are generally parameterized with some subset of observational data, and tested against remaining data. Such tests are designed to show that the model can effectively describe the observed system by demonstrating a strong correlation, or a low mean error, between prediction and observation. The danger of such a procedure is that that poorly defined parameters can be 'tuned' to give good output: when several parameters are tuned, the right answer may be generated for the wrong reason (Bagnara et al., 2014). However, it has been demonstrated than the use of data assimilation techniques, i.e. the use of several variables and data streams in the calibration process, can help greatly in reducing this problem (Williams et al., 2005)

In this study, we applied the HYDRALL model to an Italian Eddy-covariance forest site in the Italian Alps, with a Bayesian approach to calibration. The HYDRALL model (Magnani et al., 2000) is a physiology-based model, which simulates growth patterns and canopy gas exchange of even-aged coniferous stands; in particular, it contains a new representation of the effects of hydraulic constraints on carbon allocation among tree compartments. Some components of the model have been tested so far, as the response of growth and resource allocation to the environment (Magnani et

al., 2002), while the long-term dynamics of gas exchange have been compared with eddy-covariance data from pine forests across Europe (Kramer et al., 2002).

HYDRALL is a forest model that has never been calibrated with a Bayesian approach: therefore, little is known about the uncertainties associated with its predictions and its calibrated parameter values. Therefore, the aims of this work were: 1) investigate the model performances after Bayesian calibrations with increasing number of data streams ; 2) investigate the uncertainties on the model results; 3) investigate the uncertainties associated with the calibrated parameter values.

2. Materials and methods

2.1 Model Description

The HYDRALL model simulates the growth of a coniferous forest stand over a whole rotation. The model focuses on the key determinants of plant growth: light interception and gas exchange, water relations and growth allocation. In comparison with other existing forest growth models, growth allocation among tree organs is not fixed, but responds dynamically to internal and external conditions. A detailed mathematical description of the model features can be found in Magnani et al., (2002). A description of the key model features follows.

- *Light absorption and vertical profiles*

The absorption of global radiation by the sunlit and the shaded portion of a coniferous canopy, the computation of foliage isothermal net radiation and vertical integration of leaf photosynthetic parameters are based on the two-leaf model by De Pury and Farquhar (1997). A vertical exponential profile of leaf nitrogen content over the canopy is assumed, parallel to the reduction in diffuse photosynthetic photon flux density (PPFD), and both dark respiration and maximum carboxylation rates are assumed to be proportional to leaf nitrogen (Leuning, 1997). Stand aerodynamic conductance is computed iteratively as a function of wind speed and sensible heat flux from the canopy (Garratt, 1992).

- *Assimilation and respiration*

The representation of leaf assimilation is based on the Farquhar model (Farquhar and von Caemmerer, 1982) and the response of leaf assimilation to PPFD follows Farquhar and Wong (1984). The dependence of the stomatal conductance upon assimilation and air vapour pressure deficit is captured by the Leuning (1995) model, whilst a simple linear dependence of stomatal conductance upon soil water potential is assumed. It is also assumed that no gas exchange takes place whenever minimum daily air temperature falls below -5°C. Sapwood and fine root respiration are computed as a function of average daily temperature, tissue biomass and nitrogen content(Ryan, 1991). The empirical model presented by Lloyd and Taylor (1994) is used to represent the dependence of tissue (and soil) respiration upon temperature. Growth respiration is assumed to be a constant fraction of available carbon (Thornley and Johnson, 1990).

- *Water transport, growth and carbon allocation*

The transport of water through the soil–plant continuum has been modeled as described in Magnani et al. (2002). Soil water potential and hydraulic resistance are a function of soil water content, soil texture and fine root density (Campbell, 1985). Root resistance is assumed to be inversely related to fine root biomass (Magnani et al., 1996), whilst aboveground resistance is a simple function of sapwood basal area and tree height (Whitehead et al., 1984). Growth allocation among foliage, sapwood and fine roots is driven by the assumption of optimal plant growth under hydraulic constraints; a single wood compartment, comprising both stem and branches, is represented (Magnani et al., 2002, 2000). In particular, growth partitioning among tree organs is not fixed, but responds to the environment according to the hypothesis of functional homeostasis in water transport, resulting in the acclimation of the plant's structure to climatic conditions. The basic assumption is that there must be a balance in the hydraulic architecture of trees, in order to prevent xylem water potential from decreasing to values that would trigger runaway embolism and foliage dieback (Magnani et al., 2000).

- *Soil carbon dynamics*

The two-compartment model of (Andrèn and Kätterer, 1997) has been chosen to represent soil respiration and the transition from young to old soil carbon pools. A constant humification coefficient is assumed. Decomposition of young and old organic matter and humification are affected to the same extent by soil temperature and soil water potential, as captured by the multiplicative model of Andrèn and Paustian (1987).

2.2 Site description

The study site is located on the Lavarone plateau, in the western Italian Alps (Trento province, Italy). The forest is largely an even-age stand, with the oldest trees of about 130 years, dominated by Silver fir (*Abies alba* Mill., 70%) and Norway spruce (*Picea abies* (L.) Karst., 30%), with minor presence of European beech (*Fagus sylvatica* L.) and no understorey. The measurement tower is placed at the bottom of a shallow valley, at 1350 m a.s.l., closed by a 1380 m and a 1450 m high mounts to the south and north, respectively. The site characteristics of Lavarone are described in detail in Rodeghiero & Cescatti (2005).

2.3 Data

On the basis of data availability two one-year datasets were selected, one characterized by average climatic conditions for the model calibration (year 2007) and one characterized by extreme climatic conditions for the model validation (year 2003). Daily minimum, maximum and average temperatures, average relative humidity and wind speed, daily totals of global radiation and precipitation were calculated from the half hourly values provided by the PI of the site and used as input data for the model. Daily values of Net Ecosystem Productivity (NEP) and evapotranspiration (ET) were calculated from the gap-filled half hourly time series and used to calibrate the model. The nighttime measured CO_2 fluxes (thus ecosystem respiration) were corrected with a modified approach of the van Gorsel et al. (2009, 2008, 2007) method (Sottocornola et al., in prep.). This method is based on the creation of a subset of nighttime NEE, measured during the first hours of the evening, when the atmospheric conditions are still unstable. Besides for atmospheric stability, the CO_2 fluxes are excluded from this subset, if too different from the NEE modeled based on a

daytime light response (Lasslop et al., 2010). All the nighttime data are then replaced by a relationship between the monthly NEE and soil temperature means belonging on the selected subset.

The Bayesian calibration requires an estimate of the uncertainties around the data used during the calibration (van Oijen et al., 2005). Uncertainties around NEP and ET were calculated as follows:

$NEP_j = NEP_j \pm x_j$

$ET_j = ET_j \pm y_j$

where x_j and y_j are the uncertainties in day j on NEP and ET, respectively; x_j was calculated as the maximum value between 0.08 molCO$_2$ m^{-2} d^{-1} and a random value in the interval [1:0.3*NEP$_j$], while y_j was calculated as the maximum value between 0.3 mm d^{-1} and a random value in the interval [1:0.3*ET$_j$]. The lower bounds are necessary to ensure that lower values of NEP and ET would not get an overwhelming weight during the calibration procedure.

2.4 Calibration procedure

Two model calibrations were carried out: a single-variable calibration on daily NEP and a multi-variable calibration on both daily NEP and ET. For each calibration, 10 model parameters were calibrated using a Bayesian approach based on a multiple-chains MCMC sampling. To avoid the risk of providing too much *a priori* information to the model, we set the prior distributions as uniform (i.e. any value has the same probability to occur). The calibrated parameters and their meaning, together with the maximum and minimum values of the prior distributions, are listed in table 1.

For each calibration, a 3×10^5 iterations long simulation was carried out using the Differential Evolution Markov Chain (DEMC) algorithm (ter Braak, 2006). An initial burn-in phase was set to 30% of the total number of iterations. 100 chains were considered, making the number of iterations per chain equal to 3×10^3.

A sample of parameter vectors from the posterior distribution generated by the MCMC simulation was used to calculate the model results and uncertainties. The number of candidate parameter sets in the sample is equal to 1% of the total number of iterations, not considering the burn-in phase. The parameter set with maximum likelihood was used to calculate the model results, while the model uncertainties were calculated as the 5% and 95% percentile of the GPP distribution resulting from the model calculations for each parameter set in the MCMC sample.

2.5 Evaluation of model performance

Several indices were considered to evaluate model performance. A linear regression of modeled vs. measured values was fitted to calculate the coefficient of determination (R^2) and the root mean square error (RMSE) for both NEP and ET. In addition, two other tests were applied to evaluate the model performances. The first was Theil's inequality coefficient (U, Theil (1966)): Theil's U can assume values of 0 and greater. If $U = 0$ then the model produces perfect predictions. If $U = 1$ the model produces predictions of system behavior that are not better than a zero-change prediction. If $U > 1$, then the predictive power of the model is worse than the no-change prediction. The second test was the modelling efficiency (ME, Vanclay and Skovsgaard (1997)), which provides a simple index of performance on a relative scale: ME = 1 indicates a perfect fit, ME = 0 reveals that the model is no better than a simple average, while negative values indicate poor performance.

3. Results

When calibrating the model only on NEP, the model estimates of NEP were good in the calibration year, while the model strongly underestimated ET (figure 1). For both the variables, the model performed better in calibration than in validation year (table 2). The model estimated an average daily NEP of 0.21 mol CO_2 m^{-2} d^{-1} against an average measured daily NEP of 0.26 mol CO_2 m^{-2} d^{-1}, and an average daily ET of 0.51 mm d^{-1} against an average measured ET of 1.70 mm d^{-1} in the calibration year, indicating a slight underestimation of daily NEP and a very strong underestimation of daily ET. In the validation year, the model estimated an average daily NEP of

0.22 mol CO_2 m^{-2} d^{-1} against an average measured daily NEP of 0.24 mol CO_2 m^{-2} d^{-1}, and an average daily ET of 0.53 mm d^{-1} against an average measured ET of 1.81 mm d^{-1}.

When calibrating the model on both NEP and ET, the model estimates of NEP were also good in the calibration year, while the estimates of ET were only slightly worse (figure 1) and the model appeared to be able to reproduce the data. For both the variables, the model performed better in calibration than in validation year (table 2). The model estimated an average daily NEP of 0.22 mol CO_2 m^{-2} d^{-1} against an average measured daily NEP of 0.26 mol CO_2 m^{-2} d^{-1}, and an average daily ET of 1.28 mm d^{-1} against an average measured ET of 1.70 mm d^{-1} in the calibration year, indicating a slight underestimation of both daily NEP and ET. In the validation year, the model estimated an average daily NEP of 0.22 mol CO_2 m^{-2} d^{-1} against an average measured daily NEP of 0.24 mol CO_2 m^{-2} d^{-1}, and an average daily ET of 1.26 mm d^{-1} against an average measured ET of 1.81 mm d^{-1}, with an underestimation of both daily NEP and ET also in this case.

When calibrating on NEP only, the model uncertainties around the daily NEP were very small while around the daily ET were large, but nevertheless they were both lower in magnitude than the data uncertainties (figure 2). The range of uncertainties on measured NEP in 2007 varied from -0.23 to 1.05 mol CO_2 m^{-2} d^{-1}, with a maximum difference of 0.48 mol CO_2 m^{-2} d^{-1} between the daily maximum and minimum possible values, while the range of uncertainties on modeled NEP varied from -0.05 to 0.57 mol CO_2 m^{-2} d^{-1}, with a maximum difference of 0.13 molC m^{-2} d^{-1} between the daily maximum and minimum possible values. In 2003 (validation year), the range of uncertainties on measured NEP varied from -0.26 to 1.24 molC m^{-2} d^{-1}, with a maximum difference of 0.57 molC m^{-2} d^{-1} between the daily maximum and minimum possible values, while uncertainties on modeled NEP varied from -0.06 to 0.56 molC m^{-2} d^{-1}, with a maximum difference of 0.12 molC m^{-2} d^{-1} between the daily maximum and minimum possible values. Concerning ET, the range of uncertainties in 2007 varied from 0.19 to 5.85 mm d^{-1}, with a maximum difference of 2.70 mm d^{-1} between the daily maximum and minimum possible values, while the range of uncertainties on modeled ET varied from 0.00 to 2.98 mm d^{-1}, with a maximum difference of 1.60 mm d^{-1} between

the daily maximum and minimum possible values. In 2003 (validation year), the range of uncertainties on measured ET varied from 0.19 to 6.36 mm d^{-1}, with a maximum difference of 2.93 mm d^{-1} between the daily maximum and minimum possible values, while uncertainties on modeled ET varied from 0.00 to 3.15 mm d^{-1}, with a maximum difference of 1.70 mm d-1 between the daily maximum and minimum possible values.

When calibrating on both NEP and ET, model uncertainties were lower in magnitude than data uncertainties, with uncertainties on NEP higher than when calibrated on NEP only, while uncertainties on ET were greatly reduced (figure 2). The range of uncertainties on measured NEP in 2007 varied from -0.23 to 0.61 molC m^{-2} d^{-1}, with a maximum difference of 0.48 $molCO_2$ m^{-2} d^{-1} between the daily maximum and minimum possible values, while the range of uncertainties on modeled NEP varied from -0.07 to 0.61 molC m^{-2} d^{-1}, with a maximum difference of 0.35 molC m^{-2} d^{-1} between the daily maximum and minimum possible values. In 2003 (validation year), the range of uncertainties on measured NEP varied from -0.26 to 1.24 molC m^{-2} d^{-1}, with a maximum difference of 0.57 molC m^{-2} d^{-1} between the daily maximum and minimum possible values, while uncertainties on modeled NEP varied from -0.14 to 0.62 molC m^{-2} d^{-1}, with a maximum difference of 0.47 molC m^{-2} d^{-1} between the daily maximum and minimum possible values. Concerning ET, the range of uncertainties in 2007 varied from 0.19 to 5.85 mm d^{-1}, with a maximum difference of 2.70 mm d^{-1} between the daily maximum and minimum possible values, while the range of uncertainties on modeled ET varied from 0.00 to 3.85 mm d^{-1}, with a maximum difference of 2.27 mm d^{-1} between the daily maximum and minimum possible values. In 2003 (validation year), the range of uncertainties on measured ET varied from 0.19 to 6.36 mm d^{-1}, with a maximum difference of 2.93 mm d^{-1} between the daily maximum and minimum possible values, while uncertainties on modeled ET varied from 0.00 to 4.19 mm d^{-1}, with a maximum difference of 2.58 mm d^{-1} between the daily maximum and minimum possible values.

The calibrated parameter values are listed in table 3. Both calibrations narrowed the prior distributions for most of the parameters calibrated (figure 3). When using only NEP as a calibrating

variable, 7 parameters out of 10 showed narrower posterior distributions: the exceptions were the threshold of soil water potential for maximum stomatal closure (PSI0), the depth of soil explored by 90% of roots (SLDEP), and the soil water potential for null soil respiration (PSIS0). Conversely, al the parameters were constrained to peaked distributions when the model was calibrated on both NEP and ET, with PSIS0 showing the broadest posterior distribution. In general, excluding the parameters mentioned above, the posterior distributions resulting from the single-variable calibration showed lower uncertainties than the posteriors resulting from the multi-variable calibration (table 3).

4. Discussion

4.1 *Model performances*

The ability of the model to reproduce the data, hence the model performance, can be considered good in both the calibration and validation years for the variables it has been calibrated on. While the calibration on NEP only gave a good fitting on NEP but unrealistic results on ET, using both NEP and ET to calibrate the model gave good results on both variables without losing a good fitting on NEP. In both cases, the goodness of fit on NEP is comparable to the fitting obtained on NEE when applying HYDRALL on the coniferous site of Loobos (Netherlands, $R^2 = 0.67$, Kramer et al., 2002). The model was generally able to reproduce the seasonal trend of the data also during validation, with better results on ET than on NEP: this is encouraging given that 2003 can be considered an extreme year for the climatic conditions experienced by the ecosystem, with higher temperature especially during summer.

A lack of correspondence between model and data can be also partially due to uncertainties in the flux data. It has been pointed out by Raupach et al. (2005) that ''data uncertainties are as important as the data values themselves.'' The flux data may contain unsystematic and systematic errors (Aubinet et al., 2000; Wilson and Baldocchi, 2001; Wofsy et al., 1993). Furthermore, the inevitable gap-filling for missing data points in order to obtain daily values introduces additional uncertainties (Falge et al., 2001). In this particular study, given the characteristics of the site, the

total daily uncertainties on NEP an ET were both set equal to the 30% of their daily value, but the difficulties in estimating them could be partially responsible for the mismatch between modeled and measured fluxes because they can affect the effectiveness of the calibration procedure. If the data are uncertain, i.e. become less informative, then the likelihood distribution in parameter space becomes more uniform. As a consequence, every proposed new candidate parameter vector will have similar likelihood as the current parameter vector, so the likelihood ratio will always be very close to 1 and the candidate vector will always be accepted unless its prior probability is low. This very high acceptance rate will slow down the effective exploration of parameter space as the random walk loses direction, slowing down the identification of the convergence region as a direct consequence. On the other hand, if data uncertainties are too small, i.e. if the data are considered too informative, the likelihood ratio will be always close to 0, causing a very low acceptance rate. This would cause the MCMC to move very slowly through parameter space, again resulting in a delayed identification of the convergence region (Gilks et al., 1996).

4.2 *Model and parameter uncertainties*

The Bayesian approach to calibration allowed us to study both parameter and model uncertainties in detail: during the construction of the sample of the posterior parameter distribution by MCMC, the forest model needs to be run at each visited point in parameter space, so a corresponding sample of model outputs is constructed at the same time. The variation in this sample of model outputs represents the model output uncertainty. These advantages of the Bayesian calibration for process-based modeling have been outlined before by Jansen (1999) and Jansen and Hagenaars (2004) in the context of crop modeling, and by van Oijen et al. (2005) and van Oijen and Thomson (2010) in forest modelling.

Both in calibration and validation, model uncertainties around the modeled variables were lower in magnitude than data uncertainties (figure 2): although HYDRALL is a process-based model with a strong physiological background, some of its features rely on simplified mathematical descriptions of physiological processes. Moreover, no model can consider all the possible sources

of variability underlying the C and water exchanges in a complex ecosystem, which would cause the model uncertainties to be lower than data uncertainties. Some of these features are the variation of the footprint area due to wind direction and turbulence regime (Rannik et al., 2006), the variation in the contribution from trees, ground vegetation and soil to CO_2 fluxes (Davidson et al., 2006), and the error in the estimate of the total ecosystem respiration (Mäkelä et al., 2006). In contrast with Williams et al. (2005), model uncertainties were not narrowed for all the variables involved as additional data streams were incorporated in the analysis: while the uncertainties on modeled ET were greatly reduced by the incorporation of the measured ET in the calibration process, the uncertainties on modeled NEP increased in magnitude, becoming larger than the uncertainties on modeled NEP when only one variable was used in the calibration procedure. Finally, it must be also pointed out that uncertainties around model results are directly due to uncertainties around calibrated parameter values: as a result of the single-variable calibration on NEP, the uncertainties on ET were much higher than the uncertainties around the modeled ET resulting from the multi-variable calibration. This is a direct consequence of the uncertainties around the calibrated parameters SLDEP, PSI0 and PSIS0: those hydraulic parameters show well-defined peaked posterior distributions when the model is calibrated both on NEP and ET, while they show no change from the prior distributions when the model is calibrated on NEP only, and this has a direct impact on the model uncertainties in estimating ET.

The probability distribution of the parameters represents the information we have about them. Good information implies a narrow peaked probability distribution. The prior probability distribution for parameters, based only on the literature or direct measurement of the parameters, tends to be uninformative. Levy et al. (2004) searched the literature for information about the parameters of three process-based forest models (Century, BGC and Hybrid), but this led to predominantly wide, uninformative prior distributions for the parameters. Previous studies have reported limited success in estimating model parameters using eddy flux data alone: for example, Wang et al. (2007, 2001) and Knorr and Kattge (2005) found that only 3–6 parameters could be

well constrained. As a result of the Bayesian calibration, 7 out of 10 parameter distributions were narrowed when calibrating only on NEP, while all 10 parameters show narrowed peaked posterior distributions when the model is calibrated both on NEP and ET. It must be also pointed out that the calibration on the only NEP resulted in different narrowed distributions for the parameters STOM1 (coefficient in the stomatal conductance vs assimilation equation) and CSL0 (Initial organic matter in soil), with very low values of soil organic matter estimated. As pointed out by van Oijen et al. (2005), every forest model is particularly sensible to a few key output variables that, if considered for calibration purposes, might have strong impact on model parameterization and greatly improve the estimation of posterior parameter distributions. In this particular case, ET is very likely one of the key output variables for HYDRALL, while the calibration on the only NEP can be considered unreliable given the calibrated values of CSL0 and the broad posterior distributions of PSI0, SLDEP and PSIS0.

Increasing model complexity does not necessarily make for a better model (Zobitz et al., 2008), as adding on additional layers of detail may result in increased realism but also greater equifinality, and poorer performance or larger uncertainties in forward runs (e.g.,Franks and Beven, 1997), but shortening the time step at which the model runs might improve model performance (Amthor et al., 2001), and would also (1) increase the amount of eddy flux data that could be used and would eliminate the need for gap-filling; (2) improve our ability to characterize the sensitivity of fast processes to environmental drivers. Following Richardson et al. (2010), the creation of an elite dataset of daytime flux data could help a great deal in estimating nighttime values and could therefore greatly improved the model performances.

5. Conclusions

We applied the physiology-based forest model HYDRALL to an alpine coniferous forest dominated by Silver fir in the Western Italian Alps, evaluating the importance of additional information in the calibration procedure and their impact on model performances, model uncertainties, and parameter estimation. The model gave good results on NEP but poor on ET when

calibrate only on NEP, and good model performance on both NEP and ET when ET was added as a calibrating variable. Moreover, 7 parameters out of 10 showed a narrowed peaked posterior distribution when calibrating only on NEP, and all parameters showed a narrowed peaked posterior distribution when calibrating both on NEP and ET. These two results identify ET as a key output variable for HYDRALL, and prove that HYDRALL is a suitable model in this type of ecosystem. Model uncertainties were always smaller than data uncertainties, with variable magnitude depending on the calibration procedure applied, suggesting that the impact of a new variable in calibration might not help in reducing the uncertainties in modelling the other variables involved.

6. References

Amthor, J.S., Chen, J.M., Clein, J.S., Frolking, S.E., Goulden, M.L., Grant, R.F., Kimball, J.S., King, A.W., McGuire, A.D., Nikolov, N.T., Potter, C.S., Wang, S., Wofsy, S.C., 2001. Boreal forest CO2 exchange and evapotranspiration predicted by nine ecosystem process models: intermodel comparisons and relationships to field measurements. J. Geophys. Res. Atmos. 106, 33623–33648.

Andrèn, O., Kätterer, T., 1997. ICBM: the introductory carbon balance model for exploration of soil carbon balances. Ecol. Appl. 7, 1226–1236.

Andrèn, O., Paustian, K., 1987. Barley straw decomposition in the field: a comparison of models. Ecology 68, 1190–1200.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, S., Moncrieff, J., Foken, T., Kowalski, A.S., Martin, P.H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J.A., Granier, A., Grünwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., Vesa, T., 2000. Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. Adv. Ecol. Res. 30, 113–175.

Bagnara, M., Sottocornola, M., Cescatti, A., Minerbi, S., Montagnani, L., Gianelle, D., Magnani, F., 2014. Bayesian optimization of a light use efficiency model for the estimation of daily gross primary productivity in a range of Italian forest ecosystems. Ecol. Modell. doi:10.1016/j.ecolmodel.2014.09.021

Baldocchi, D., Meyers, T., 1998. On using eco-physiological , micrometeorological and biogeochemical theory to evaluate carbon dioxide , water vapor and trace gas fluxes over vegetation : a perspective. Agric. For. Meteorol. 1–25.

Campbell, G.S., 1985. Soil Physics with BASIC Transport Models for Soil–Plant Systems. Elsevier, Amsterdam.

Canadell, J.G., Le Quéré, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A., Marland, G., 2007. Contributions to accelerating atmospheric

CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. Proc. Natl. Acad. Sci. 104 , 18866–18870. doi:10.1073/pnas.0702737104

Davidson, E.A., Richardson, A.D., Savage, K.E., Hollinger, D.Y., 2006. A distinct seasonal pattern of the ratio of soil respiration to total ecosystem respiration in a spruce-dominated forest. Glob. Chang. Biol. 12, 230–239. doi:0.1111/j.1365-2486.2005.01062.x

De Pury, D.G.G., Farquhar, G.D., 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant, Cell Environ. 20, 537–557. doi:10.1111/j.1365-3040.1997.00094.x

Ellison, A.M., 2004. Bayesian inference in ecology. Ecol. Lett. 7, 509–520.

Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecol. Appl. 6, 1036–1046.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C.T., Law, B.E., Meyers, T., Moncrieff, J., Moors, E., Munger, J.W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. For. Meteorol. 107, 43–69. doi:10.1016/S0168-1923(00)00225-2

Farquhar, G.D., von Caemmerer, S., 1982. Modelling of photosynthetic response to environmental conditions., in: Lange, O.L., Nobel, P.S., Osmond, C.B., Ziegler, H. (Eds.), Physiological Plant Ecology II. Springer Verlag, Berlin, pp. 549–587.

Farquhar, G.D., Wong, S.C., 1984. An empirical model of stomatal conductance. Aust. J. Plant Physiol. 11, 191–210.

Field, C.B., 2001. Plant Physiology of the "Missing" Carbon Sink. Plant Physiol. 125, 25–28.

Franks, S.W., Beven, K.J., 1997. Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. J. Geophys. Res. 102, 23,991–23,999.

Friend, A.D., Schugart, H.H., Running, S.W., 1993. A Physiology-Based Gap Model of Forest Dynamics Ecology Vol. 74, No. 3 (Apr., 1993), pp. 792-797. Ecology 74, 792–797.

Garratt, J.R., 1992. The Atmospheric Boundary Layer. Cambridge University Press, Cambridge, UK.

Ghazoul, J., McAllister, M., 2003. Communicating complexity and uncertainty in decision-making contexts: Bayesian approaches to forest research. Int. For. Rev. 5, 9–19.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov chain Monte Carlo in practice. CRC press.

Goodale, C.L., Apps, M.J., Birdsey, R.A., Field, C.B., Heath, L.S., Houghton, R.A., Jenkins, J.C., Kohlmaier, G.H., Kurz, W., Liu, S., Nabuurs, G.-J., Nilsson, S., Shvidenko, A.Z., 2002. Forest

Carbon Sinks In The Northern Hemisphere. Ecol. Appl. 12, 891–899. doi:10.1890/1051-0761(2002)012[0891:FCSITN]2.0.CO;2

Jansen, M.J.W., 1999. Data use and Bayesian statistics for model calibration., in: Stein, A., Penning de Vries, F.W.T.. (Eds.), Data and Models in Action. Kluwer, Dordrecht, pp. 69–80.

Jansen, M.J.W., Hagenaars, T.J., 2004. Calibration in a Bayesian modelling framework., in: Van Boekel, M.A.J.S., Stein, A., Van Bruggen, A.H.C. (Eds.), Bayesian Statistics and Quality Modelling in the Agro-Food Production Chain. Kluwer, Dordrecht, pp. 47–55.

Knorr, W., Kattge, J., 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Glob. Chang. Biol. 11, 1333–1351.

Kramer, K., Leinonen, I., Bartelink, H.H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, a. J., Froer, O., Gracia, C. a., Granier, a., Grunwald, T., Hari, P., Jans, W., Kellomaki, S., Loustau, D., Magnani, F., Markkanen, T., Matteucci, G., Mohren, G.M.J., Moors, E., Nissinen, a., Peltola, H., Sabate, S., Sanchez, a., Sontag, M., Valentini, R., Vesala, T., 2002. Evaluation of six process-based forest growth models using eddy-covariance measurements of CO2 and H2O fluxes at six forest sites in Europe. Glob. Chang. Biol. 8, 213–230. doi:10.1046/j.1365-2486.2002.00471.x

Landsberg, J.J., Gower, S.T., 1997. Applications of physiological ecology to forest management. Academic Press: London.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A.D., Arneth, A., Barr, A., Stoy, P., Wohlfahrt, G., 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Glob. Chang. Biol. 16, 187–208. doi:10.1111/j.1365-2486.2009.02041.x

Leuning, R., 1997. Scaling to a common temperature improves the correlation between the photosynthesis parameters Jmax and Vcmax. J. Exp. Bot. 48, 345–347.

Leuning, R., 1995. A critical appraisal of a combined stomatalphotosynthesis model for C3 plants. Plant, Cell Environ. 18, 339–355.

Levy, P.E., Wendler, R., van Oijen, M., Cannell, M.G.R., Millard, P., 2004. The effects of nitrogen enrichment on the carbon sink in coniferous forests: uncertainty and sensitivity analyses of three ecosystem models. Water, Air, Soil Pollut. 4, 67–74.

Lloyd, J., Taylor, J.A., 1994. On the Temperature Dependence of Soil Respiration. Funct. Ecol. 8, 315–323.

Magnani, F., Centritto, M., Grace, J., 1996. Measurement of apoplasmic and cell-to-cell components of root hydraulic conductance by a pressure clamp technique. Planta 199, 296–306.

Magnani, F., Grace, J., Borghetti, M., 2002. Adjustment of tree structure in response to the environment under hydraulic constraints. Funct. Ecol. 16, 385–393.

Magnani, F., Mencuccini, M., Grace, J., 2000. Age-related decline in stand productivity: the role of structural acclimation under hydraulic constraints. Plant, Cell Environ. 23, 251–263. doi:10.1046/j.1365-3040.2000.00537.x

Mäkelä, A., Kolari, P., Karimäki, J., Nikinmaa, E., Perämäki, M., Hari, P., 2006. Modelling five years of weather-driven variation of GPP in a boreal forest. Agric. For. Meteorol. 139, 382–398. doi:0.1016/j.agrformet.2006.08.017

Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., Liski, J., Zhou, L., Alexeyev, V., Hughes, M.K., 2001. A large carbon sink in the woody biomass of Northern forests. Proc. Natl. Acad. Sci. 98 , 14784–14789. doi:10.1073/pnas.261555198

Rannik, Ü., Kolari, P., Vesala, T., Hari, P., 2006. Uncertainties in measurement and modelling of net ecosystem exchange of a forest. Agric. For. Meteorol. 138, 244–257. doi:10.1016/j.agrformet.2006.05.007

Raupach, M.R., Rayner, P.J., Barrett, D.J., DeFries, R.S., Heimann, M., Ojima, D.S., Quegan, S., Schmullius, C.C., 2005. Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Glob. Chang. Biol. 11, 378–397. doi:10.1111/j.1365-2486.2005.00917.x

Reckhow, K.H., 1990. Bayesian inference in non-replicated ecological studies. Ecology 71, 2053–2059.

Richardson, A.D., Williams, M., Hollinger, D.Y., Moore, D.J.P., Dail, D.B., Davidson, E.A., Scott, N.A., Evans, R.S., Hughes, H., Lee, J.T., Rodrigues, C., Savage, K., 2010. Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints. Oecologia 164, 25–40. doi:10.1007/s00442-010-1628-y

Rodeghiero, M., Cescatti, A., 2005. Main determinants of forest soil respiration along an elevation/temperature gradient in the Italian Alps. Glob. Chang. Biol. 11, 1024–1041. doi:10.1111/j.1365-2486.2005.00963.x

Ryan, M.G., 1991. A simple method for estimating gross carbon budgets for vegetation in forest ecosystems. Tree Physiol. 9, 255–266.

Sivia, D.S., 1996. Data Analysis: A Bayesian Tutorial. Oxford University Press, Oxford.

Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., Lindroth, a, 2008. Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand. Ecol. Modell. 213, 331–344. doi:10.1016/j.ecolmodel.2008.01.001

Ter Braak, C.J.F., 2006. A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. Stat. Comput. 16, 239–249. doi:10.1007/ s11222-006-8769-1

Theil, H., 1966. Applied Econometric Forecasting. North-Holland, Amsterdam.

Thornley, J.H.M., Johnson, I.R., 1990. Plant and Crop Modeling. Oxford University Press.

Van Gorsel, E., Delpierre, N., Leuning, R., Black, A., Munger, J.W., Wofsy, S., Aubinet, M., Feigenwinter, C., Beringer, J., Bonal, D., Chen, B., Chen, J., Clement, R., Davis, K.J., Desai, A.R., Dragoni, D., Etzold, S., Gu, L., Heinesch, B., Hutyra, L.R., Jans, W.W.P., Kutsch, W., Law, B.E., Leclerc, M.Y., Mammarella, I., Montagnani, L., Noormets, A., Rebmann, C., Wharton, S., 2009. Agricultural and Forest Meteorology Estimating nocturnal ecosystem respiration from the vertical turbulent flux and change in storage of CO 2. Agric. For. Meteorol. 149, 1919–1930.

Van Gorsel, E., Leuning, R., Cleugh, H., Keith, H., Suni, T., 2007. Nocturnal carbon efflux: reconciliation of eddy covariance and chamber measurements using an alternative to the u*-threshold filtering technique. Tellus B 59, 397–403. doi:10.1111/j.1600-0889.2007.00252.x

Van Gorsel, E., Leuning, R., Cleugh, H.A., Keith, H., Kirschbaum, M.U.F., Suni, T., 2008. Application of an alternative method to derive reliable estimates of nighttime respiration from eddy covariance measurements in moderately complex topography. Agric. For. Meteorol. 148, 1174–1180. doi:10.1016/j.agrformet.2008.01.015

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, 915–27. doi:10.1093/treephys/25.7.915

Van Oijen, M., Thomson, 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. Clim. Change 103, 55–67. doi:10.1007/s10584-010-9917-3

Vanclay, J.K., Skovsgaard, J.P., 1997. Evaluating forest growth models. Ecol. Modell. 98, 1–12. doi:10.1016/S0304-3800(96)01932-1

Wang, Y.P., Baldocchi, D.D., Leuning, R., Falge, E., Vesala, T., 2007. Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites. Glob. Chang. Biol. 13, 652–670.

Wang, Y.P., Leuning, R., Cleugh, H.A., Coppin, P.A., 2001. Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful? Glob. Chang. Biol. 7, 495–510.

White, J.., Coops, N.., Scott, N.., 2000. Estimates of New Zealand forest and scrub biomass from the 3-PG model. Ecol. Modell. 131, 175–190. doi:10.1016/S0304-3800(00)00251-9

Whitehead, D., Edwards, W.R.N., Jarvis, P.G., 1984. Conducting sapwood area, foliage area, and permeability in mature trees of Picea sitchensis and Pinus contorta. Can. J. For. Res. 14, 940–947.

Williams, M., Schwarz, P. a., Law, B.E., Irvine, J., Kurpius, M.R., 2005. An improved analysis of forest carbon dynamics using data assimilation. Glob. Chang. Biol. 11, 89–105.

Wilson, K.B., Baldocchi, D.D., 2001. Comparing independent estimates of carbon dioxide exchange over 5 years at a deciduous forest in the southeastern United States. J. Geophys. Res. 106, 34167–34178. doi:10.1029/2001JD000624

Wofsy, S.C., Goulden, M.L., Munger, J.W., Fan, S.-M., Bakwin, P.S., Daube, B.C., Bassow, S.L., Bazzaz, F.A., 1993. Net Exchange of CO2 in a Mid-Latitude Forest. Sci. 260 , 1314–1317. doi:10.1126/science.260.5112.1314

Zobitz, J.M., Moore, D.J.P., Sacks, W.J., Monson, R.K., Bowling, D.R., Schimel, D.S., 2008. Integration of process-based soil respiration models with whole-ecosystem CO2 measurements. Ecosystems 11, 250–269.

7. Tables and figures

Table 1: calibrated parameters in the HYDRALL model, and their minimum and maximum values in the prior distributions. Ds: vapour pressure deficit; gs: stomatal conductance; A: assimilation rate.

Parameter	Meaning	Unit	Minimum	Maximum
CSL0	Initial organic matter in soil	$kgDM\ m^{-2}$	0	50
PSI0	Threshold of soil water potential for maximum stomatal closure	MPa	-10	-10^{-5}
SLDEP	Depth of soil explored by 90% of roots	M	10^{-5}	4
STOM1	Coeff. in gs vs A equation	Pa	10^{-5}	3×10^{6}
TCHIL	Maximum temperature inducing chilling	°C	-5	10
VCMPK	Peak seasonal value of maximum carboxylation rate at optimum temperature	$mol\ m^{-2}\ s^{-1}$	10^{-6}	10^{-3}
VPD0	Coeff. in gs response to Ds	Pa	10^{-5}	10^{4}
WFUND	Understorey foliage biomass	$kgDM\ m^{-2}$	10^{-8}	0.1
PSIS0	Soil water potential for null soil respiration	MPa	-15	-0.1
Q10_SAT	Respiration Q10 at soil water capacity	-	0.5	4

Table 2: Model performances in both calibration and validation (R^2: coefficient of determination; RMSE: root mean square error; U: Theil's inequality coefficient; ME: modelling efficiency) using only NEP (s) or both NEP and ET (m) as calibrating variables.

	Variable	R^2 (s)	RMSE (s)	U (s)	ME (s)	R^2 (m)	RMSE (m)	U (m)	ME (m)
Calibration	ET-2007	0.60	1.38	0.71	0.18	0.53	0.76	0.39	0.45
	NEP-2007	0.65	0.15	0.42	0.62	0.63	0.15	0.42	0.61
Validation	ET-2003	0.52	1.54	0.41	0.48	0.55	0.91	0.73	0.17
	NEP-2003	0.47	0.19	0.49	0.57	0.44	0.19	0.53	0.46

Table 3: Calibrated parameter values (as the parameter vector with the maximum log-likelihood) and their uncertainties (as the standard error of the posterior distribution for each parameter, excluding the burn-in phase) using only NEP (s) or both NEP and ET (m) as calibrating variables.

Parameter	Calibrated value (s)	Calibrated value (m)
CSL0	5.27	29.02
(se)	(1.09×10^{-2})	(1.67×10^{-2})
PSI0	-1.22	-0.49
(se)	(6.31×10^{-3})	(5.08×10^{-3})
SLDEP	0.06	0.23
(se)	(2.57×10^{-3})	(2.12×10^{-3})
STOM1	1.40×10^{5}	2.998×10^{6}
(se)	(1.03×10^{3})	(1.69×10^{3})
TCHIL	2.86	0.79
(se)	(2.39×10^{-3})	(4.41×10^{-3})
VCMPK	1.15×10^{-4}	5.57×10^{-5}
(se)	(1.65×10^{-7})	(3.38×10^{-7})
VPD0	9.95×10^{3}	167.60
(se)	(4.67)	(3.98)
WFUND	0.097	0.097
(se)	(1.98×10^{-5})	(4.24×10^{-5})
PSIS0	-1.90	-12.66
(se)	(9.53×10^{-3})	(8.21×10^{-3})
Q10_SAT	3.02	3.99
(se)	(1.42×10^{-3})	(1.03×10^{-3})

Figure 1. Measured and modeled time series of NEP and ET (triangles and dashed line: data; dots and continuous line: model estimates)
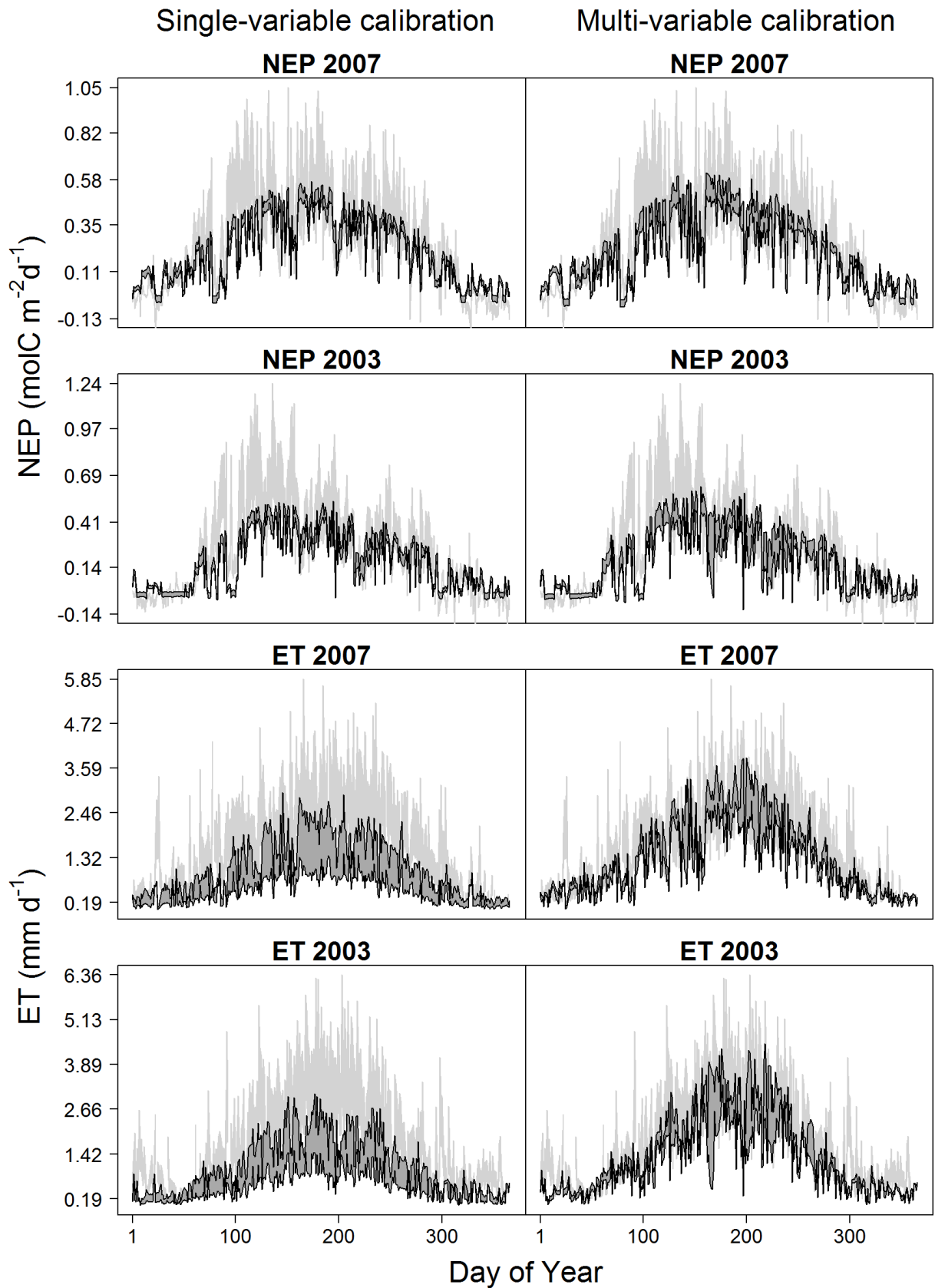
Figure 2. Uncertainties around measured (light-grey) and modeled (dark-grey) NEP and ET.
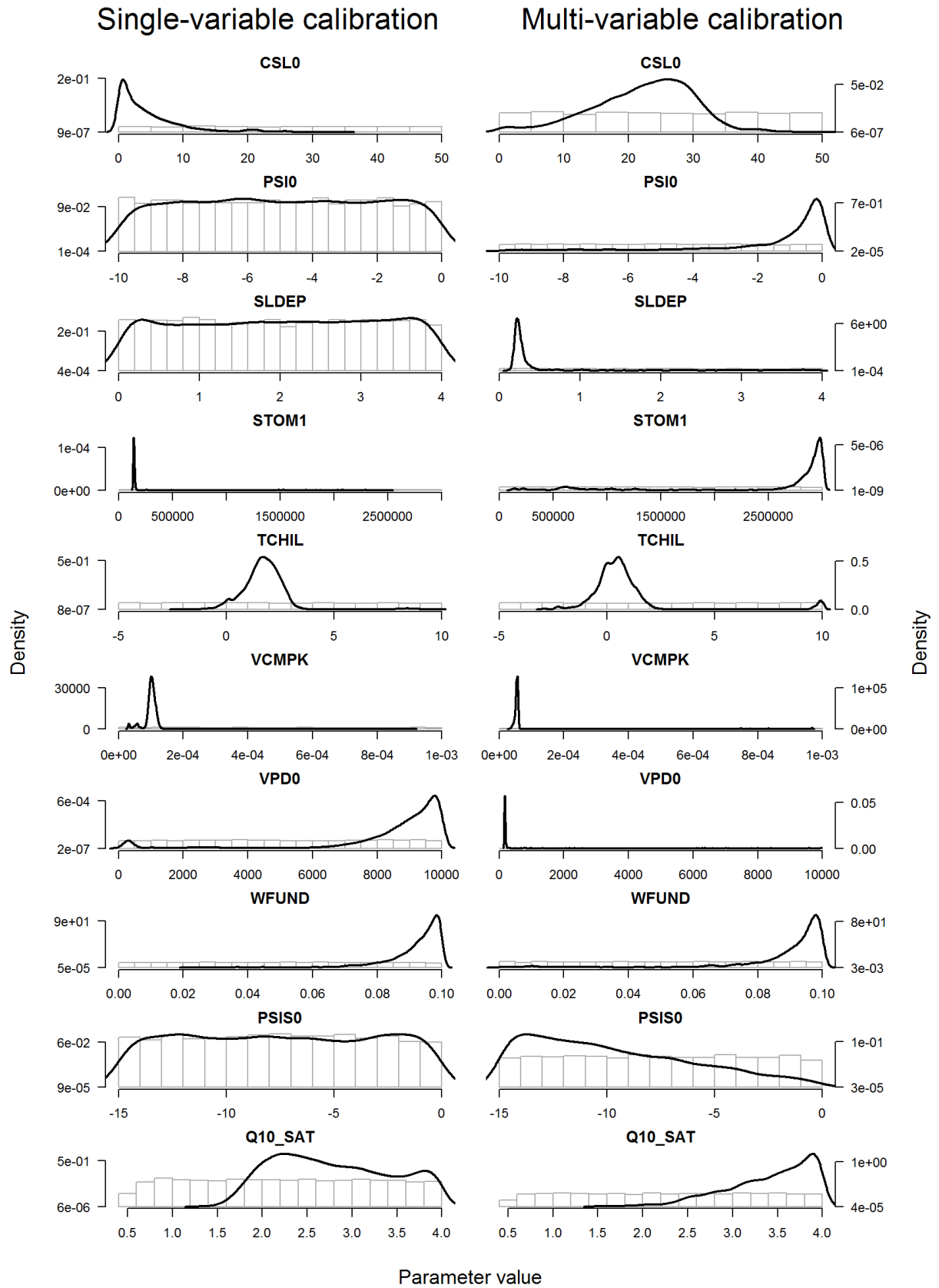
Figure 3. Prior (histograms) and posterior (lines) distributions for the 10 calibrated parameters.

Chapter 5: Conclusions

This study aimed at testing the Bayesian procedure for calibration to different types of forest models, to evaluate their performances and the uncertainties associated with them. In particular, the main objectives of this thesis were the following:

- to apply a Bayesian framework to calibrate forest models and test their performances in different biomes and different environmental conditions;

- to identify and solve structure-related issues in simple models;

- to identify the advantages of additional information made available when calibrating forest models with a Bayesian approach.

The first objective was investigated in Chapter 2, with the application of the Prelued model to eight Italian eddy-covariance sites. The ability of Prelued to reproduce the estimated Gross Primary Productivity (GPP) was tested over contrasting natural vegetation types that represented a wide range of climatic and environmental conditions. This simple light-use efficiency (LUE)-based model proved to be versatile, performing well in almost all sites considered with the exception of a Mediterranean macchia, suggesting that it can be successfully applied to a wider range of forest types than previously tested. However, some issues related to Prelued's multiplicative mathematical structure were highlighted, given the model unrealistic responses to driving climatic variables in three Mediterranean sites(e.g. model fitting in three Mediterranean sites). The Bayesian approach to calibration did not bring a clear advantage concerning model performances, yet it allowed to estimate uncertainties around all parameter values, highlighting some possible convergence issue during the calibration process.

The issues related to Prelued's multiplicative structure were the main topic of Chapter 3: focusing on the forest site of Lavarone (Trento province, Italy) as a case study, three different MCMC-based algorithms were applied within a Bayesian framework to calibrate the model, and their performances were compared. The application of the three different algorithms of increasing complexity (Metropolis-Hastings Random Walk, Adaptive Metropolis, Differential Evolution

Markov Chain) with different number of iterations showed that all three MCMC-methods were similarly effective in reaching convergence. For all of them a very high number of iterations (10^6) was required for the Markov Chain to stabilize in the convergence region. This was due to the combination of at least two different factors: 1) a strongly multiplicative mathematical structure, and 2) empirical parameters with neither ecological nor physiological meaning about which little is known in the literature. To reduce the computational effort to reach proper convergence, a two-steps calibration proved to be an adequate method, using the posterior correlations between parameters resulting from the first step to reduce the number of parameters calibrated in the second step.

A different, more complex model was applied in Chapter 4, which focused on the application of the physiology-based forest model HYDRALL to the forest ecosystem of Lavarone. This exercise aimed at evaluating the importance of additional information in the calibration procedure and their impact on model performances, model uncertainties, and parameter estimation. The model gave good results on Net Ecosystem Productivity (NEP) but poor on evapotranspiration (ET) when calibrated only on NEP, and good model performance on both NEP and ET when ET was added as a calibrating variable. Moreover, 7 parameters out of 10 showed a narrowed peaked posterior distribution when calibrating only on NEP, while all 10 parameters showed a narrowed peaked posterior distribution when calibrating both on NEP and ET. These two results identified ET as a key output variable for HYDRALL, and proved that HYDRALL is a suitable model for this type of ecosystem. Model uncertainties were always smaller than data uncertainties, with different magnitude depending on the calibration procedure applied, suggesting that the impact of a new variable in calibration might not help in reducing the uncertainties in modelling the other variables involved.

Overall, the Bayesian technique proved to be an excellent and versatile tool to successfully calibrate forest models of different structure and complexity, on different kind and number of variables and with a different number of parameters involved. However, given its relatively recent history as a tool in forest modelling, several further steps can be recommended:

1) The application of Prelued to several Eddy-Covariance sites highlighted the necessity of an improvement of the estimates of uncertainties around daily GPP derived from eddy-covariance measurements, together with more information on the prior probability distributions of the model parameters.

2) Even when working with simple models, there is the necessity to always investigate the model behavior from the ecological point of view in order to identify possible issues in the calibration procedure. Those issues can be solved, as shown in Chapter 3, with the application of different Monte Carlo Markov Chain algorithms, but it should also be pointed out that LUE-based models do not allow to compare model estimates against actual data: GPP is not measured, it is derived from NEP. Therefore, NEP should be the model output against which the calibration should be performed; thus it should be included in LUE models via combination with a respiration model.

3) Another important point relates to the empirical nature of the parameters: when possible, a large use of parameters with no physical or physiological meaning should be avoided, in order to rely on the physiological basis of GPP as much as possible. When doing so, as was done in Chapter 4, the Bayesian calibration of even more complex models was not as problematic as in the former cases. Nevertheless the results of the calibration on NEP alone showed some issues: 3 parameters out of 10 were not well defined and daily ET was strongly underestimated. Therefore, the identification of key model outputs is paramount for the success of such a calibration, and the addition of those outputs to the calibration process is likely to improve the model performances, reduce the model uncertainties, and allow for a better estimation of parameter values and of their probability distributions.

4) Especially for process-based models, shortening the time step at which the model runs: a) might improve model performance, b) would increase the amount of data that could be used, c) would eliminate the need for gap-filling, and d) would improve our ability to characterize the sensitivity of fast processes to environmental drivers.

Acknowledgments