Ecological Modelling 341 (2016) 37–52

Contents lists available at ScienceDirect

Ecological Modelling

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

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

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

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

ARTICLE INFO

*Article history:*
Received 16 December 2015
Received in revised form
20 September 2016
Accepted 22 September 2016
Available online 4 October 2016

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

ABSTRACT

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

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

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



## 1. Introduction

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

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

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

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

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

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

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

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

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

## 2. Materials and methods

### 2.1. PRELES model

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

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

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

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

### 2.2. Carbon and water flux data

Stand-scale net ecosystem exchange (NEE) of $CO_2$, evapotranspiration and meteorological data from ten boreal coniferous forest sites located in Finland and Sweden were used in this study (Table 1). The sites cover a latitudinal band from 60°N to 67°N with annual mean temperatures ranging from 0.8 to 7.1 °C, and precipitation from ~550 mm to ~850 mm. Leaf area index (LAI) at each site was treated as one lumped LAI, i.e. all the canopy layers were included in one unique layer. The total (all-sided) LAI varies between ~3.8 and ~12 m$^2$ m$^{-2}$ offering a good possibility to address both climatic and LAI controls on forest GPP and ET. A brief summary of the sites is provided in Table 1, and complete descriptions can be found in the respective References

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