



# Land-surface parameter optimisation using data assimilation techniques: the adJULES system V1.0

Nina M. Raoult, Tim E. Jupp, Peter M. Cox, and Catherine M. Luke

National Centre for Earth Observation, University of Exeter, Exeter EX4 4QF, UK

Correspondence to: Nina M. Raoult (nr278@exeter.ac.uk)

Received: 29 December 2015 – Published in Geosci. Model Dev. Discuss.: 20 January 2016

Revised: 11 July 2016 – Accepted: 19 July 2016 – Published: 25 August 2016

**Abstract.** Land-surface models (LSMs) are crucial components of the Earth system models (ESMs) that are used to make coupled climate–carbon cycle projections for the 21st century. The Joint UK Land Environment Simulator (JULES) is the land-surface model used in the climate and weather forecast models of the UK Met Office. JULES is also extensively used offline as a land-surface impacts tool, forced with climatologies into the future. In this study, JULES is automatically differentiated with respect to JULES parameters using commercial software from FastOpt, resulting in an analytical gradient, or adjoint, of the model. Using this adjoint, the adJULES parameter estimation system has been developed to search for locally optimum parameters by calibrating against observations. This paper describes adJULES in a data assimilation framework and demonstrates its ability to improve the model–data fit using eddy-covariance measurements of gross primary production (GPP) and latent heat (LE) fluxes. adJULES also has the ability to calibrate over multiple sites simultaneously. This feature is used to define new optimised parameter values for the five plant functional types (PFTs) in JULES. The optimised PFT-specific parameters improve the performance of JULES at over 85 % of the sites used in the study, at both the calibration and evaluation stages. The new improved parameters for JULES are presented along with the associated uncertainties for each parameter.

on providing the lower boundary condition for atmospheric models by calculating the land–atmosphere fluxes of heat, moisture, and momentum, and updating the surface state variables on which these fluxes depend (e.g. soil temperature, soil moisture, snow cover). In the mid- to late 1990s some land-surface modelling groups began to introduce additional aspects of biology into their schemes, most notably the dynamic control of transpiration by leaf stomata and the connected rates of leaf photosynthesis (Sellers et al., 1997; Cox et al., 1999).

In the early 2000s, climate modelling groups began to use the carbon fluxes simulated by LSMs within first generation climate–carbon cycle models (Cox et al., 2000; Friedlingstein et al., 2001). These early results, and a subsequent model inter-comparison (Friedlingstein et al., 2006), highlighted the uncertainties associated with land carbon–climate feedbacks. The 5th Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR5; Stocker et al., 2013) for the first time routinely included models with an interactive carbon cycle (now called Earth system models or ESMs), confirming that land responses to climate and CO<sub>2</sub> are amongst the largest of the uncertainties in future climate change projections (Arora and Boer, 2005; Brovkin et al., 2013; Jones et al., 2013; Friedlingstein et al., 2013). Any future decreased ability of the land surface to drawdown atmospheric CO<sub>2</sub> could imply smaller “compatible emissions” in order to stay below key warming thresholds such as 2 °.

Uncertainties in LSMs arise from three major sources: parameter uncertainty, process uncertainty, and uncertainty due to initial and boundary conditions. Taking these in reverse order, uncertainty due to initial and boundary conditions refers to uncertainty in the forcing data (Kavetski et al., 2006a, b; Ajami et al., 2007). Process uncertainty includes the misrep-

## 1 Introduction

Land-surface models (LSMs) have formed an important component of climate models for many decades now (Pitman, 2003). First generation land-surface schemes focussed

resentation of land-surface processes and also the neglect of important processes (such as nitrogen limitations on plant growth, see for example Thornton et al., 2007; Zaehle et al., 2010), or canopy light interception (Mercado et al., 2009). The drive to reduce process uncertainty almost invariably leads to increases in LSM complexity, which typically leads to the introduction of additional internal model parameters. Parameter uncertainty arises from uncertainty in these internal model parameters. The evolution of LSMs has therefore involved an attempt to reduce process uncertainty by increasing model realism/complexity, but at the cost of increasing parameter uncertainty. This paper concerns the development and application of a technique to reduce parameter uncertainty in the widely used Joint UK Land Environment Simulator (JULES) LSM (Best et al., 2011; Clark et al., 2011).

Optimisation techniques come under the umbrella of model–data fusion and range from simple ad hoc parameter tuning to rigorous data assimilation frameworks. These approaches have been used in a number of studies, covering various LSMs, to derive vectors of parameters that improve model–data fit significantly (e.g. Wang et al., 2001, 2007; Reichstein et al., 2003; Knorr and Kattge, 2005; Raupach et al., 2005; Santaren et al., 2007; Thum et al., 2008; Williams et al., 2009; Peng et al., 2011). Many of these studies calibrate the model at individual measurement sites. Given the small spatial footprint of each flux tower, this can often result in over-tuning. This over-tuning may occur when a single site does not represent the full range of a plant functional type (PFT), given different tree types, tree ages, and above-ground biomass found at each site. There may be some anomalous plants in the small footprint that are not representative of the PFTs over a broader area. The optimised model parameters are site specific and often struggle to perform as well when generalised over other sites (Xiao et al., 2011).

The majority of LSMs group vegetation into a small number of PFTs. Model parameters are assumed to be generic over each PFT. Through different optimisation techniques, some studies have tried to assess the robustness of PFT-specific parameters (e.g. Kuppel et al., 2014). Medvigy et al. (2009) and Verbeeck et al. (2011) both showed that parameters derived at one site can perform well on a similar site and over the surrounding region (Medvigy and Moorcroft, 2011). However, a contradictory study by Groenendijk et al. (2010) found that there was cross-site parameter variability after optimisation within the PFT groupings.

In the last few years, there has been a move towards deriving PFT-specific parameters using data from multiple sites, the results of which have been generally positive (e.g. Xiao et al., 2011 and Kuppel et al., 2012). Both of these studies used data from multiple sites in their optimisation (calling it multi-site optimisation) and have commented on the robustness of this technique, showing that the choice of the initial parameter vector had little effect on the optimised values.

Kuppel et al. (2012) compared different approaches for finding generic PFT-specific parameters, such as averaging

optimised parameter vectors over PFTs and directly optimising over multiple sites. They found that the latter method was best for finding PFT-specific parameters. The multi-site optimisation procedure was refined in Kuppel et al. (2014), extended to other PFTs, and evaluated at a global scale.

For global modelling, there is a clear need to find generic parameters and associated uncertainties for each PFT, by optimising against observations in a reproducible way. This paper presents a model–data fusion framework, called adJULES, that allows data from multiple sites to be used simultaneously in order to improve the JULES land-surface model. The adJULES system uses the adjoint method, which finds minima rapidly across multiple parameters via matrix inversion and has the advantage of reproducibility. Replicating these findings using brute-force optimisation would be prohibitively expensive computationally.

This paper aims to answer the following questions:

- Can a (locally) optimum vector of generic parameters for each of the JULES PFT classes be found?
- How does the optimal PFT parameter vector compare to parameter vectors found by optimising each site individually?
- How robust is the adJULES system when optimising over multiple sites?
- What uncertainty is associated with each parameter?

In Sect. 2, methods and data used in the study are described. The JULES land-surface model and our new data assimilation system (adJULES), are introduced, along with the data used and the parameters chosen to be optimised in the study. In Sect. 3, the results are presented. The methodology for optimising over multiple sites simultaneously is validated, and optimum parameter values are provided for each JULES PFT. The performance of the new parameter sets is assessed and shown to significantly improve the fit of the JULES model to the observations. The conclusions are laid out in Sect. 4.

## 2 Methods and data

### 2.1 The JULES land-surface model

The JULES land-surface model (Best et al., 2011; Clark et al., 2011) simulates the interactions between the land and the atmosphere. Originally developed from the Met Office Surface Exchange Scheme (MOSES) (Cox et al., 1999), JULES can be used “offline” with observed atmospheric forcing data, or can be coupled into a global circulation model (GCM). JULES is the land-surface model used in the UK Met Office Unified Model.

JULES is a mechanistic land-surface model including physical, biophysical, and biochemical processes that control the radiation, heat, water, and carbon fluxes in response

**Table 1.** Parameters in optimisation vector, with descriptions.

Symbol	Name in code	Description	Units
$n_0$	n10	Top leaf nitrogen concentration	kg N (kg C) <sup>-1</sup>
$f_0$	f0	Maximum ratio of internal to external CO <sub>2</sub>	–
$d_r$	rootd_ft	Root depth	m
$\alpha$	alpha	Quantum efficiency	mol CO <sub>2</sub> per mol PAR photons
$\frac{\delta c}{\delta l}$	dcatch_dlai	Rate of change of canopy interception capacity with LAI (leaf area index)	kg m <sup>-2</sup>
$T_{low}$	tlow	Lower temperature for photosynthesis	°C
$T_{upp}$	tupp	Upper temperature for photosynthesis	°C
$dq_c$	dqcrit	Humidity deficit at which stomata close	kg kg <sup>-1</sup>

to time series of the state of the overlying atmosphere (Best et al., 2011; Clark et al., 2011). Processes such as evaporation, plant growth, and soil microbial activity are all linked through mathematical equations that quantify how environmental conditions affect evapotranspiration, heat balance, respiration, photosynthesis, and carbon assimilation (Best et al., 2011; Clark et al., 2011). JULES runs at a given sub-daily step (typically 30 min), using meteorological drivers of rainfall, incoming radiation, temperature, humidity, and wind speed as inputs.

Vegetation in the JULES model is categorised into five PFTs: broadleaf trees (BT), needleleaf trees (NT), C3 grasses (C3G), C4 grasses (C4G), and shrubs (Sh). Default parameters for these PFT classes are taken from a previous study (Blyth et al., 2010).

The eight parameters that are calibrated within this study (see Table 1) relate predominantly to leaf-level stomatal conductance ( $g$ ) and photosynthesis ( $A$ ). Four of these parameters control the responses of  $g$  and  $A$  to environmental conditions such as surface temperature ( $T_{upp}$ ,  $T_{low}$ ), solar radiation ( $\alpha$ ), and atmospheric humidity deficit ( $dq_c$ ). Two other calibration parameters ( $f_0$ ,  $n_{l0}$ ) essentially control the maximum values of  $g$  and  $A$ . The remaining two calibration parameters influence the hydrological partitioning at the land surface and relate to the amount of rainfall intercepted by the plant canopy ( $\delta c/\delta L$ ), and the “root depth” ( $d_r$ ) from which each PFT can access soil water for transpiration. The simulated latent heat flux and gross primary productivity have been found to be especially sensitive to these parameters in previous studies (Blyth et al., 2010).

The full set of equations within the JULES model is documented in the literature by Best et al. (2011) and Clark et al. (2011), but the key equations are highlighted below. In JULES, leaf-level photosynthesis and stomatal conductance are treated with a coupled model (Cox et al., 1998). Based on the models of Collatz et al. (1991, 1992), leaf-level photosynthesis  $A$  is controlled by the carboxylation rate (which depends on  $n_0$ ,  $T_{low}$ ,  $T_{upp}$ ) and light-limited photosynthesis (which depends on  $\alpha$ ). It follows that

$$A = A(n_0, \alpha, T_{low}, T_{upp}, c_i, \beta), \tag{1}$$

where  $c_i$  is the internal CO<sub>2</sub> concentration inside the leaf, and  $\beta$  is a soil moisture stress factor, which depends on the vertical soil moisture profile  $\theta$ , and the plant root depth  $d_r$ :

$$\beta = \beta(\theta, d_r). \tag{2}$$

The internal CO<sub>2</sub> concentration  $c_i$  is assumed to be dependent on the external CO<sub>2</sub> concentration  $c_a$  and the atmospheric humidity deficit  $\delta q$  (Cox et al., 1998) via the equation

$$\frac{c_i - c_*}{c_a - c_*} = f_0 \left( 1 - \frac{\delta q}{\delta q_c} \right), \tag{3}$$

where  $c_*$  is the CO<sub>2</sub> compensation point, and  $f_0$  and  $\delta q_c$  are parameters that are calibrated in this study. The stomatal conductance for water vapour  $g$  is diagnosed in JULES from the leaf-level photosynthesis  $A$  and the internal and external CO<sub>2</sub> concentrations:

$$g = 1.6 \frac{A}{c_a - c_i}. \tag{4}$$

The factor of 1.6 converts the stomatal conductance for CO<sub>2</sub> into a stomatal conductance for water vapour. The scaling-up from leaf to canopy level in this version of JULES uses a “big-leaf” approach (Cox et al., 1999).

## 2.2 Data assimilation system

The term “data assimilation” is commonly used to describe the process of using observations to refine the initial state within a numerical representation of a system (Bouttier and Courtier, 1999). This is most obviously the case for weather forecasting, in which the temperature, humidity, and wind fields define the initial state. However, data assimilation techniques have also been used for parameter estimation, for example in hydrological models (Madsen, 2003; Liu and Gupta, 2007), and carbon cycle data assimilation systems (CCDAS; Rayner et al., 2005; Kaminski et al., 2013). In parameter optimisation by data assimilation, the internal parameters of a model take on the role of the dynamical state variables in initial state estimation by data assimilation. Nevertheless, the underlying techniques (e.g. of defining a model adjoint and minimising the error in the fit to data), are very

similar in these two applications of data assimilation. This paper is certainly not the first to define parameter estimation of this form as data assimilation (Braswell et al., 2005; Stöckli et al., 2008; Verbeeck et al., 2011; Kuppel et al., 2012; Hararuk et al., 2014), but the reader should note the subtle difference between our definition of data assimilation and that commonly used in weather forecasting.

Even a relatively simplistic land-surface representation such as JULES has over a hundred internal parameters representing the environmental sensitivities of the various land-surface types and PFTs within the model. In general these parameters are chosen to represent measurable “real-world” quantities (e.g. aerodynamic roughness length, surface albedo, plant root depth). This allows for observationally based estimates of these parameters to be made in the early stages of the model development process. However, the detailed performance of a land-surface model can be very sensitive to such internal parameters. It is therefore common for land-surface modellers to calibrate their models against available observations, such as eddy-covariance flux data. This is typically carried out in a rather ad hoc manner with the modeller varying the parameters that he/she believes are most relevant to the model performance. Such model tuning is by its very nature subjective, lacks reproducibility, and is often sub-optimal because the modeller is unable to explore the full feasible parameter space through such a manual technique.

This paper describes a more objective approach to land-surface model calibration, adopting ideas from the applied mathematics of data assimilation as used widely in weather forecasting and other disciplines, and motivated by pioneering attempts at carbon cycle data assimilation (Rayner et al., 2005; Kaminski et al., 2013). It utilises the adjoint of the JULES model, derived by automatic differentiation, which enables efficient and objective calibration against observations. Importantly, adJULES also allows the uncertainties in the best-fit parameters to be estimated. Such uncertainties are important information for model users, and can also form the basis for observation-constrained estimates of posterior probability density functions for the land-surface parameter perturbations used in climate model ensembles (e.g. Booth et al., 2012).

### 2.2.1 The theory of adJULES

JULES generates a modelled time series for a given vector of internal parameters,  $\mathbf{z}$ . The cost function,  $f(\mathbf{z})$ , consists of a weighted sum of squares of the difference between  $\mathbf{m}_t$  (the vector of model outputs at time  $t$ ), and  $\mathbf{o}_t$  (the vector of observations at time  $t$ ), combined with a term quadratic in the difference between parameter values  $\mathbf{z}$  and initial parameter values  $\mathbf{z}_0$ :

$$f(\mathbf{z}; \hat{\mathbf{z}}, \mathbf{z}_0) = \frac{1}{2} \left[ \sum_t (\mathbf{m}_t(\mathbf{z}) - \mathbf{o}_t)^T \mathbf{R}(\hat{\mathbf{z}})^{-1} (\mathbf{m}_t(\mathbf{z}) - \mathbf{o}_t) + \lambda (\mathbf{z} - \mathbf{z}_0)^T \mathbf{B}^{-1} (\mathbf{z} - \mathbf{z}_0) \right]. \quad (5)$$

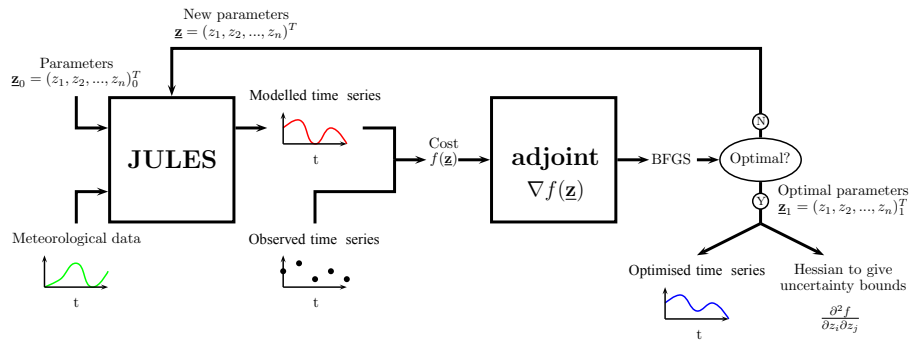
Here,  $\mathbf{R}(\hat{\mathbf{z}}) = \frac{1}{n} \sum_{t=1}^n (\mathbf{m}(\hat{\mathbf{z}})_t - \mathbf{o}_t)(\mathbf{m}(\hat{\mathbf{z}})_t - \mathbf{o}_t)^T$  denotes the error cross-product matrix produced by a JULES run with a parameter value  $\hat{\mathbf{z}}$ . In an optimisation,  $\mathbf{z}$  and  $\hat{\mathbf{z}}$  are updated separately in nested loops, having both been initialised to the default JULES parameter value  $\mathbf{z}_0$ . In the inner loop,  $\mathbf{z}$  is varied to minimise the cost function (termination criterion:  $\nabla f \approx 0$ ) for the current value of  $\hat{\mathbf{z}}$ . In the outer loop,  $\hat{\mathbf{z}}$  is reset to the new value of  $\mathbf{z}$  from the inner loop (termination criterion: change in  $\hat{\mathbf{z}}$  negligible). At the end of an optimisation, therefore, the matrix  $\mathbf{R}$  conveys information about the error correlation structure in a JULES run with optimal parameter values.

The matrix  $\mathbf{B}$  describes the prior covariances assigned to the parameters, and is here chosen to be a diagonal matrix proportional to the inverse square of the ranges allowed for each parameter. The prior uncertainties are therefore assumed to be uncorrelated between the parameters. The constant of proportionality  $\lambda$  controls the relative importance of the background term (i.e. the right-hand term in Eq. 5) and the error term (i.e. the left-hand term in Eq. 5). Larger values of  $\lambda$  help condition the problem and force parameter values to be close to the initial value  $\mathbf{z}_0$  (Bouttier and Courtier, 1999). All parameters and observations are equally weighted in this cost function.

The optimal vector of parameters is the vector  $\mathbf{z}$  that minimises the cost function (Eq. 5). The aim of adJULES is to find this vector. adJULES minimises the cost function iteratively using the gradient descent algorithm L-BFGS-B (Byrd et al., 1995, optim: R Development Core Team, 2015). This algorithm is based on the BFGS quasi-Newton method but is modified to use limited memory, for computational affordability, and box constraints, so each parameter is given an upper and lower bound based on expert opinion or on physical reasoning (Byrd et al., 1995).

At each iteration, the gradient  $\nabla f(\mathbf{z})$  of the cost function  $f(\mathbf{z})$  is computed with respect to all parameters, using the adjoint model of JULES. The adjoint is generated with the automatic differentiator tool TAF (Transformation of Algorithms in Fortran; see Giering et al., 2005). Automatic differentiation relies on using the chain rule; the choice of forward or reverse mode refers to the order in which the derivatives are computed. Calculating  $\nabla f(\mathbf{z})$  is most efficient in reverse mode as only one sweep is needed to generate the derivative with respect to all parameters (Bartholomew-Biggs et al., 2000).

Once the cost function reaches the minimum, a locally optimal parameter vector  $\mathbf{z}_1$  is returned and the second derivative of the cost function with respect to the parameters can be used to calculate posterior uncertainties. This process is



**Figure 1.** Schematic of the adJULES parameter estimation system starting with the initial parameter vector  $z_0$ . This is usually based on default JULES parameter values (Blyth et al., 2010). The optimised parameter vector is denoted  $z_1$ .

then repeated, the locally optimised parameters are fed back through JULES, generating a new modelled time series and hence a new cost function. The loop terminates when the modelled time series no longer improves (Fig. 1).

### 2.2.2 Multi-site implementation

In its simplest form, adJULES runs at a single grid-point location and so the derived optimal parameter vector is site specific. On the other hand, multi-site optimisation aims to find values for a common set of parameters, using data from multiple locations. The definition of the cost function (Eq. 5) can be extended to include the observations from all  $S$  sites, and its derivative found in order to use the L-BFGS-B algorithm again. The extended cost function is the sum of the individual cost functions for each site  $s$ . Similarly, the first and second derivatives of this new cost function can be defined using the sum of the derivatives at the individual sites.

$$f(\mathbf{z}; \hat{\mathbf{z}}, \mathbf{z}_0) = \frac{1}{2} \left[ \sum_s \sum_t (\mathbf{m}_{t,s}(\mathbf{z}) - \mathbf{o}_{t,s})^T \mathbf{R}_s(\hat{\mathbf{z}})^{-1} (\mathbf{m}_{t,s}(\mathbf{z}) - \mathbf{o}_{t,s}) + S\lambda(\mathbf{z} - \mathbf{z}_0)^T \mathbf{B}^{-1}(\mathbf{z} - \mathbf{z}_0) \right] \quad (6)$$

An additive cost function, where the optimisation criterion is to minimise the total cost, was chosen over a cost function where all individual cost functions are required to be improved. All of the sites were used in finding the optimal parameter vector for each PFT, so that sites that do not improve with the rest of the PFT suggest incorrect classification of the site or issues with the PFT definitions.

### 2.3 Eddy-covariance flux data

The eddy-covariance flux data used in this study are part of FLUXNET (Baldocchi et al., 2001). The FLUXNET database contains more than 500 locations worldwide, and all of the data are processed in a harmonised manner using the standard methodologies including correction, gap-filling, and partitioning (Papale et al., 2006). Data from 160 sites

were made available for this study by M. Groenendijk. The sites used in this study were selected based on data availability: sites with missing input variables or data gaps of more than 50 % during the growing season were omitted.

To constrain photosynthetic parameters, net ecosystem exchange (NEE) and latent heat (LE) flux, among other fluxes, are helpful. The NEE flux, defined as the net flux of  $\text{CO}_2$ , is partitioned into gross primary production (GPP) and ecosystem respiration (Resp) (Reichstein et al., 2005). In this study the GPP flux is used along with the LE flux to constrain the model. GPP data are model-derived estimates, which could introduce an additional uncertainty into the results. Also note that due to model structural errors, calibration against these two particular observables could cause model simulations of other fluxes (not used in the tuning) to become worse (Gupta et al., 1999).

In an attempt to run the experiments as closely to a standard JULES run as possible, input fields of vegetation structure and soil type were drawn from the UK Met Office ancillary files used in the HadGEM2 configurations. The LAI (leaf area index) seasonal cycle used is derived from a MODIS product (Myneni et al., 2002) from Boston University. The values taken for each of the experiment sites correspond to the closest grid point at which data are available. This could lead to inconsistencies between the actual vegetation at a given site and the vegetation structure and soil type used in the model.

### 2.4 Experimental set-up

Version 2.2 of JULES is implemented in the current version of adJULES. This version is set up to calibrate a subset of JULES soil and vegetation parameters against up to six observables in the vectors  $\mathbf{m}_t$  and  $\mathbf{o}_t$  (Eq. 5): NEE, sensible heat ( $H$ ), LE, surface temperature ( $T_s$ ), GPP, and Resp.

This study aims to improve the parameters used to define PFTs and therefore it concentrates on vegetation parameters. Table 1 outlines the parameters chosen. Note that this only represents a modest subset of the parameters available and as

such the results could be different when considering different subsets in the calibration.

In all, 1 year of FLUXNET data is used for each site considered in this study at the calibration stage. Where multiple years are available, the most complete year was chosen. For each site the model is spun up to a steady soil moisture and temperature state. Where possible, the 2 years of data preceding the year of comparison were applied repeatedly in the spin up. Where this was not possible, the first year of data was repeatedly applied. Only sites with at least 2 years of data are used in this study, so that the spin-up year is different from the experiment year. In each case, the model was spun up for at least 50 years. For deciduous sites and crop sites, leaf area index values are taken from MODIS data for the appropriate year. Where possible, a second year of FLUXNET data was spun up to be used at the evaluation stage of this study. This second year was chosen to be the second most complete year when more than 1 year was available.

The sites used in each of the PFT classes are described in Appendix A1. The FLUXNET database used in this study did not distinguish between the different types of grasslands. Using Met Office ancillary files, the grasslands were partitioned into C3 grasses and C4 grasses according to fractional cover. In the case of C3 grasses, sites were picked only when the fractional cover was over 60 %. Since the C4 grasses are under-represented in the FLUXNET database, this boundary was lowered to include all sites where C4 grass was the dominant PFT. Crops were not included in either grass class. The photosynthesis model used in JULES is based on scaling up observed processes at the leaf scale to represent the canopy. The scaling to canopy level can be done in several ways. In this study the simple big leaf approach was adopted (Clark et al., 2011), although optimisations can also be carried out for more complex canopy radiation options (Mercado et al., 2009).

All of the sites in each PFT class are used to find the optimal values for the PFT. The second derivative of the cost function found by differentiation of the adjoint code is then used to quantify the uncertainties associated with these optimal parameter vectors.

Preliminary experiments showed very narrow uncertainties whilst running the optimisation scheme over multiple sites (i.e. the background term was found to dominate the cost function). In previous multi-site studies (Kuppel et al., 2012, 2014), the prior range was also used to defined the background covariance matrix **B**. The range was variously further multiplied by a factor of 40 % (Kuppel et al., 2012) and one-sixth (Kuppel et al., 2014). Experiments were run to find a similar factor to use in this study (the constant of proportionality  $\lambda$  in Eq. 5). In each of the multi-site experiments, the lowest value of  $\lambda$  such that the Hessian is positive definite at the optimal parameter value was used. This allows uncertainties to be generated around each parameter and prevents the gradient descent algorithm from reaching the boundaries of the prescribed prior range.

## 2.5 Analysis tools

### 2.5.1 Parameter uncertainty

As well as generating optimal parameter values, adJULES estimates the uncertainty associated with each parameter. The second derivative (Hessian) of the cost function,

$$H_{ij} = \frac{\partial^2 f}{\partial z_i \partial z_j}, \quad (7)$$

where  $f(z)$  is given by Eq. (5), evaluated at the optimal parameter value, yields information about the curvature of the cost function at the local minimum. A “sharp” cost function, where the cost function is steep either side of the optimal parameter value, indicates lower parameter uncertainty. This can also be interpreted as meaning that a small deviation from the optimal parameter value yields a large increase in cost. Conversely, a “flat” cost function indicates higher parameter uncertainty, or little change in cost caused by deviation from the optimal parameter value.

In order to generate statistics associated with the curvature of the cost function, the Hessian is used to generate samples from the posterior distribution. This is a truncated multi-variate normal distribution (Genz et al., 2015) because of the box constraints placed on the prior. Using Gibbs sampling (Geman and Geman, 1984), an ensemble of plausible parameter vectors is generated from this distribution, for a statistically satisfactory match between observations and modelled time series. The multi-variate normal parameter distribution allows for marginal density plots to be generated for each parameter. When considering these marginal density plots, it is important to remember that they represent only one or two dimensions of a high-dimensional multi-variate normal distribution which is truncated. Consequently, the optimal parameter values (which are modes of the full high-dimensional distribution) may not coincide with modes of the one- and two-dimensional (2-D) marginal distributions.

In order to illustrate the parameter uncertainties, error bars are used to represent the 80 % quantile range (10th to 90th percentile) for each optimal parameter.

### 2.5.2 Fractional error – a metric of model–data fit

To measure the improvement exhibited by different parameter vectors, the fraction of variance unexplained  $\epsilon^2$  is used to define the fractional error  $\epsilon$ . This metric was chosen to show not only the improvement made by the optimal parameter vectors at each site but also how each site performed relative to others.

Given a parameter vector,  $z$ , a modelled time-series  $m_{i,t}$  with  $k$  data points is generated using JULES, where  $i$  denotes one of the observable data streams (in this case LE and GPP). For each data stream  $i$ , the fraction of variance unexplained by the model is

$$\epsilon_i^2 = \frac{\sum_{t=1}^k (o_{i,t} - m_{i,t})^2}{\sum_{t=1}^k (o_{i,t} - \bar{o}_i)^2}, \text{ where } \bar{o}_i = \frac{1}{k} \sum_{t=1}^k o_{i,t}. \quad (8)$$

It follows that the mean fraction of variance unexplained across data streams,

$$\epsilon^2 = \frac{\epsilon_1^2 + \epsilon_2^2}{2}, \quad (9)$$

is a single dimensionless measure of model misfit. The fractional error  $\epsilon$  can then be interpreted as the typical (root-mean-square) error expressed as a fraction of the (root-mean-square) magnitude of the observed seasonal cycle. Thus,  $\epsilon = 0$  represents a perfect match to the observations, while  $\epsilon = 1$  corresponds to the error in a null model whose prediction  $m_{i,t}$  always equals the observational mean  $\bar{o}_i$ .

In hydrology this is related to a metric known as the ‘‘Nash–Sutcliffe’’ efficiency (Nash and Sutcliffe, 1970), equivalent to  $1 - \epsilon^2$ , and has been used by many studies to perform cross-site comparisons.

### 3 Results and discussion

In this section, the site-specific optimisations are considered first. By considering each PFT separately, the misfits between the model and the observations are discussed and the effect of optimising over each site individually to improve model–observation agreement is considered.

Next, the multi-site methodology is used to perform optimisations over each of the PFTs. All of the sites in a given PFT are optimised simultaneously to find a generic parameter vector appropriate to the PFT. The new optimised parameter vectors are presented, along with associated uncertainties. Some of the uncertainties and correlations found between parameters are discussed, especially in the context of the equations described in Sect. 2.1. The rest of the section considers the improvement found using these optimised parameter vectors both on the calibration year and the evaluation year for each of the sites.

#### 3.1 Single-site optimisations

First, each of the sites was optimised individually in order to find site-specific parameter vectors. Typically, this required about 150 function evaluations to find a local optimum. As described in Sect. 2.4, 1-year runs at the different sites were optimised against monthly averaged LE and GPP. The constant of proportionality  $\lambda$  is set to 1 for all sites, in order to give equal weighting to both terms in Eq. (5). A site dominated by each PFT was picked to represent the general improvements made. The main seasonal cycles of LE and GPP for the different sites are shown in Fig. 2.

Most broadleaf sites follow the pattern illustrated (Fig. 2, top row). Normally, for broadleaf sites, a standard JULES run

will underestimate GPP. The optimisation does a good job in correcting this, bringing the modelled time series closer to the observations. In contrast, LE does not improve as much.

Similarly for the needleleaf sites (Fig. 2, second row), the JULES model output tends to overestimate LE and underestimate GPP. The parameter vector found in the optimisation improves the fit of both data streams, most notably GPP. At sites for which a double peak seasonality is apparent, the optimised model captures this better than the original model.

GPP is also underestimated for the C3 grass sites (Fig. 2, middle row) and, for the majority of the sites, the optimisation does a good job of correcting this. The LE flux tends to have the right magnitude before optimisation, unlike the GPP flux, but adJULES does not manage to improve this output significantly. In the example shown, the JULES model using the default parameter vector already performs very well, so little improvement is needed, but this is not always the case. The new set of parameters is also good at simulating multiple peaks in the LE and GPP fluxes, when they are observed.

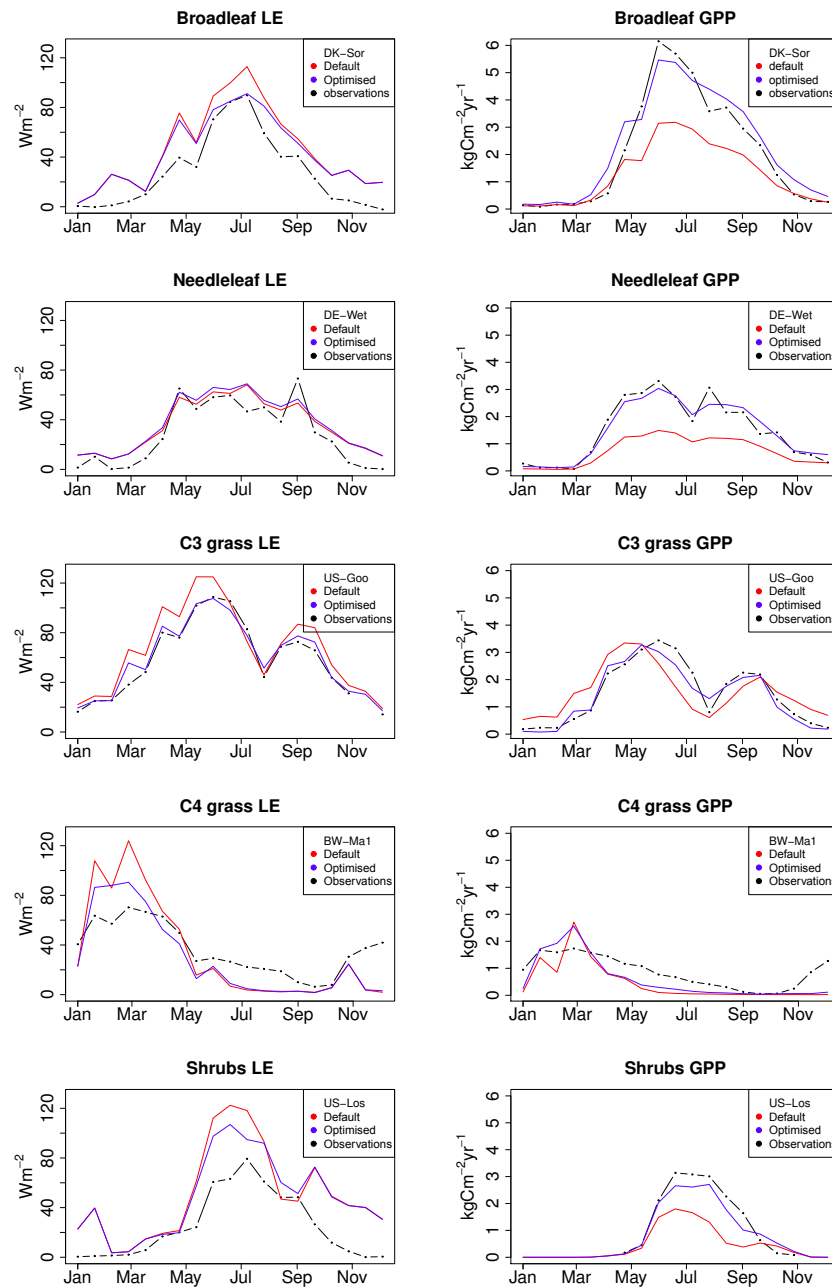
There are only two C4 grass sites in the set and JULES does not perform very well on these before or after optimisation (Fig. 2, fourth row). The original stomatal conductance–photosynthesis model within JULES was developed based on fluxes measured over C4 grass as part of the FIFE field experiment (Cox et al., 1998). However, there are relatively few FLUXNET sites over C4-dominated landscapes, and only two in the extended data set used here. As a result, the sensitivity of stomatal conductance and photosynthesis to environmental factors has been less well tested for C4 grasses. The results presented in this paper therefore highlight the need to reassess JULES and other land-surface models for predominantly C4 landscapes.

The shrub sites show no general pattern (Fig. 2, fourth row). Some sites overestimate LE, whilst others underestimate it, and similarly for GPP. The level of improvement varies over sites. For some of the sites in this PFT, the magnitude of GPP fails to get close to the magnitude of the observations, both before and after optimisation. However, it is hard to pick out a general pattern for this PFT, since there are only five sites in this set.

Overall, the adJULES system works well in finding optimal parameter vectors, which improve the performance of JULES at individual sites, regardless of PFT. The systematic underestimation of GPP in default JULES improves the most. This larger improvement in the GPP fit reflects the larger set of optimised parameters that are exclusively related to the carbon cycle. Different parameters may need to be incorporated, for example some soil ones, for the LE flux to improve further.

#### 3.2 PFT-specific optimal parameter values

Optimisations were performed over all available sites for each of the PFTs simultaneously. The optimised model parameters for each of the PFTs are presented in Fig. 3.



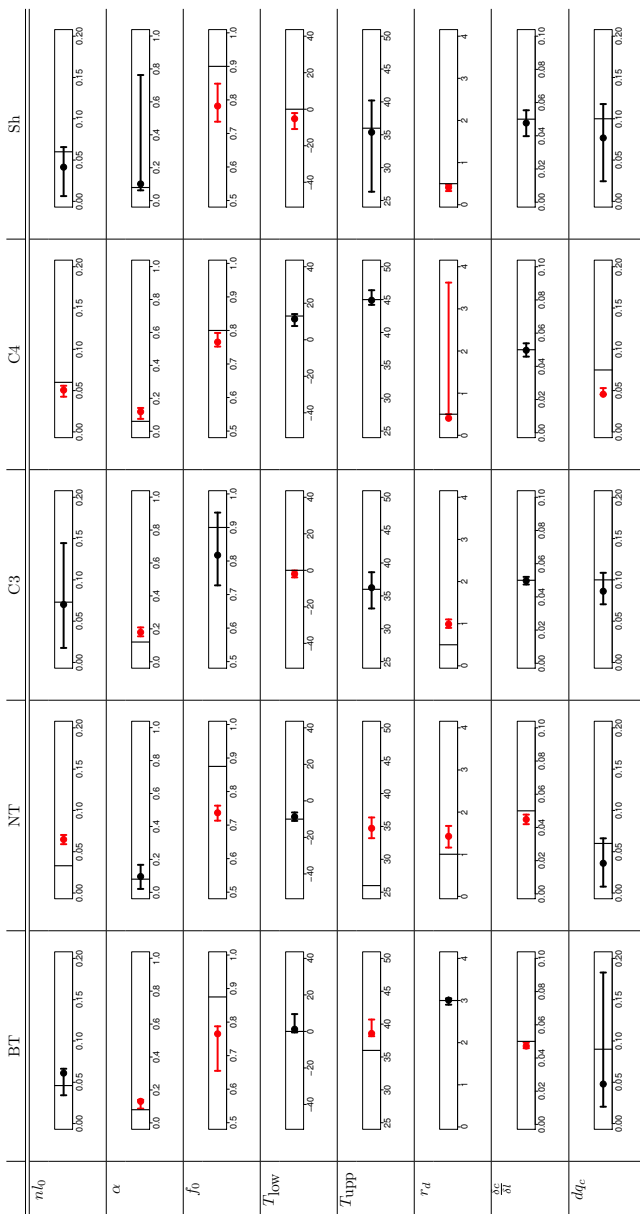
**Figure 2.** Time-series plots for illustrative site-specific evaluations showing LE (left) and GPP (right) for each of the different PFTs. Observations (black) are compared to JULES runs using default parameters (red) and site-specific optimal parameters (blue).

For half of the parameters, the prior parameter value lies outside the posterior uncertainty bounds. The  $\frac{\delta c}{\delta l}$  parameter, which determines the efficiency of rainfall interception by the plant canopy, does not change much from its original value for any of the PFTs. The uncertainty bounds are relatively tight and symmetrical. The rest of the parameters show more variation. As described in Sect. 2.5.1, the optimal values need not be in the centre of the uncertainty range, the PDF (probability density function) can be skewed. Most of the PFTs display high uncertainty in at least one of the pa-

rameters optimised; for the optimised broadleaf set for example,  $dq_c$  is highly unconstrained. For C4 grasses,  $d_r$  is so unconstrained that the optimal value found lies outside the 80 % confidence interval. C3 grasses show large uncertainty in  $n_0$  and, for shrubs, the parameter with the largest uncertainty is  $\alpha$ .

The uncertainties shown in Fig. 3 are 1-D marginal distributions. To understand further how the parameters are correlated, consider the 2-D representation in Fig. 4. For all of the PFTs, the posterior parameter uncertainties exclude a large





**Figure 3.** Summary of PFT-specific optimal JULES parameters found in this study (Table 1). The error bars show the uncertainty ranges given as an 80% quantile interval. The range of each box is the prior range of the parameters. Highlighted in red are the error bars for which the prior values (vertical line) are found outside the posterior uncertainty bounds. A numerical version of this figure is given in Table B1.

part of the prior ranges. The cloud of plausible points tends to be restrictive and tight for most parameters.

Figure 4 shows clear correlation of some parameters, especially for the tree PFTs. Many of these correlations can be understood in terms of the underlying structure of the JULES model (Sect. 2.1). For example, the positive correlation of  $n_0$  with  $f_0$ , and the negative correlation of  $n_0$  with  $dq_c$ , are

consistent with adJULES attempting to fit the stomatal conductance  $g$ , which controls the transpiration flux from taller vegetation. The stomatal conductance has the approximate form

$$g \approx 1.6 \frac{A}{c_a} \left( \frac{1}{(1 - f_0) + f_0 \frac{dq}{dq_c}} \right) \quad (10)$$

if it is assumed that  $c_* \ll c_i$  and  $c_* \ll c_a$ .

The maximum rate of leaf photosynthesis is controlled largely by the leaf nitrogen content  $n_0$ , especially in this big-leaf version of JULES (Cox et al., 1999). The best-fit parameters for tree PFTs also seem to imply that the second term in the denominator dominates over the first. As a result, maintaining a realistic  $g$  value, and therefore a realistic LE flux, will require that  $n_0$  and  $f_0$  vary proportionally, and that  $n_0$  and  $dq_c$  values are negatively correlated. This is consistent with Fig. 4a and b.

Such correlation of parameters is less obvious for the grass PFTs, because evapotranspiration is controlled less by stomatal conductance and more by the smaller aerodynamic conductances associated with shorter vegetation.

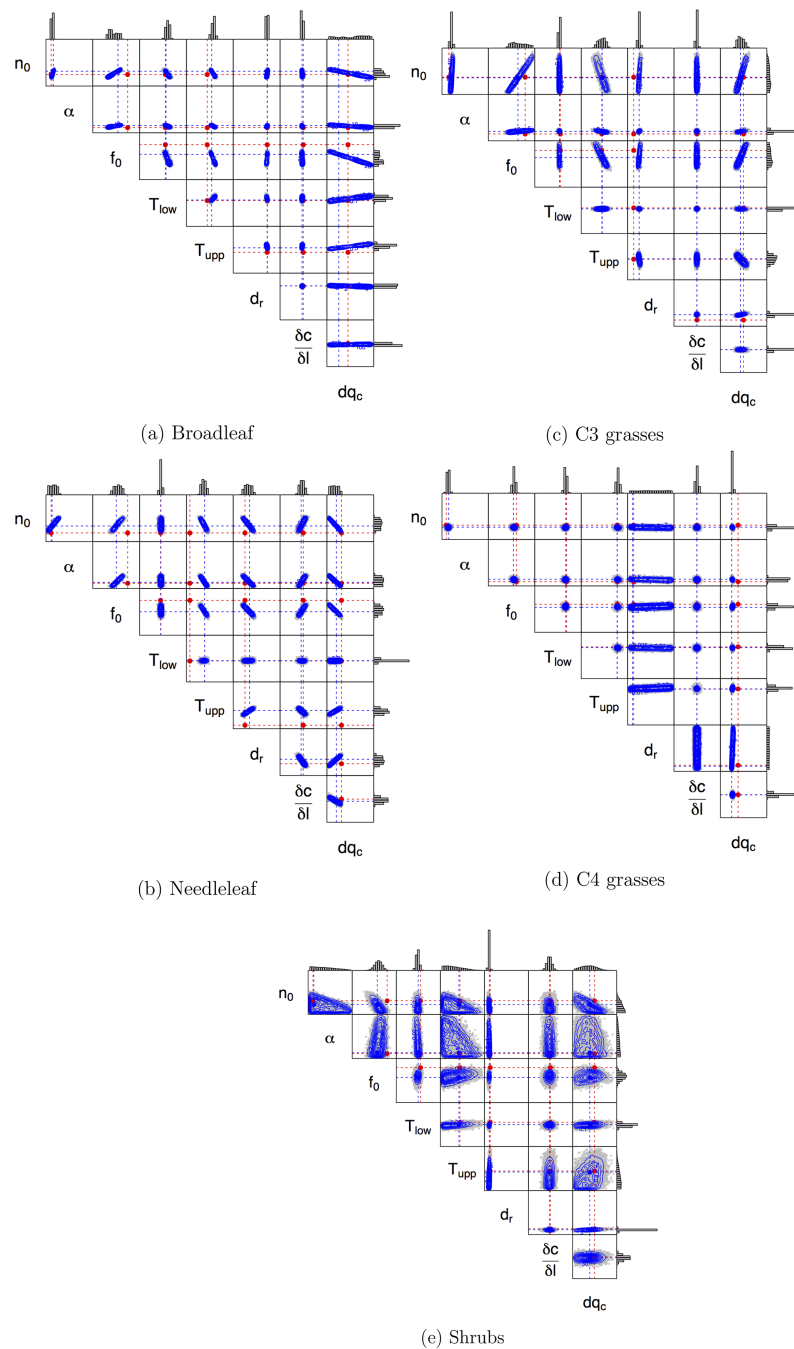
Choice of  $\lambda$  had less effect on the values of the optimal parameters than on uncertainties and correlations found. The uncertainty ranges become larger with smaller values  $\lambda$  and correlations less pronounced.

### 3.3 Assessment of PFT-specific optimal parameters

The performance of the PFT-specific parameters is now compared to the default JULES values and to the parameters found by optimising independently at each measurement site. For each site, the fractional error in both the calibration year and the evaluation year is displayed Fig. 5.

By definition, the fractional error in calibration years decreases when moving from default to site-specific optimal parameters in the calibration years. Remarkably, the site-specific optimal parameters also improve the model–data fit in evaluation years for 59/64 (92%) of sites. Similarly, the PFT-specific optimal parameter vector improves the fit (in both calibration and evaluation years) for 85% of the sites; 75/79 sites for the calibration years and 55/64 sites for the evaluation years.

Consider first the broadleaf sites (Fig. 5, top two rows). For the majority of sites displayed in the top broadleaf panel, the reduction in fractional error in moving from default to site-specific optimal parameters is substantial and sometimes as much as a factor of 2. In the calibration year, the PFT-specific optimal parameter vector improves 26 of the 27 broadleaf sites shown although one of the sites, IT-Lec, the fit shows no change. The improvement is typically about half as good (on a log scale) as the improvement using the site-specific optimal parameters. In other words, the reduction in fractional error moving from default to PFT-specific optimal parameters is sometimes as much as a factor of  $\sqrt{2}$ . Amongst broadleaf

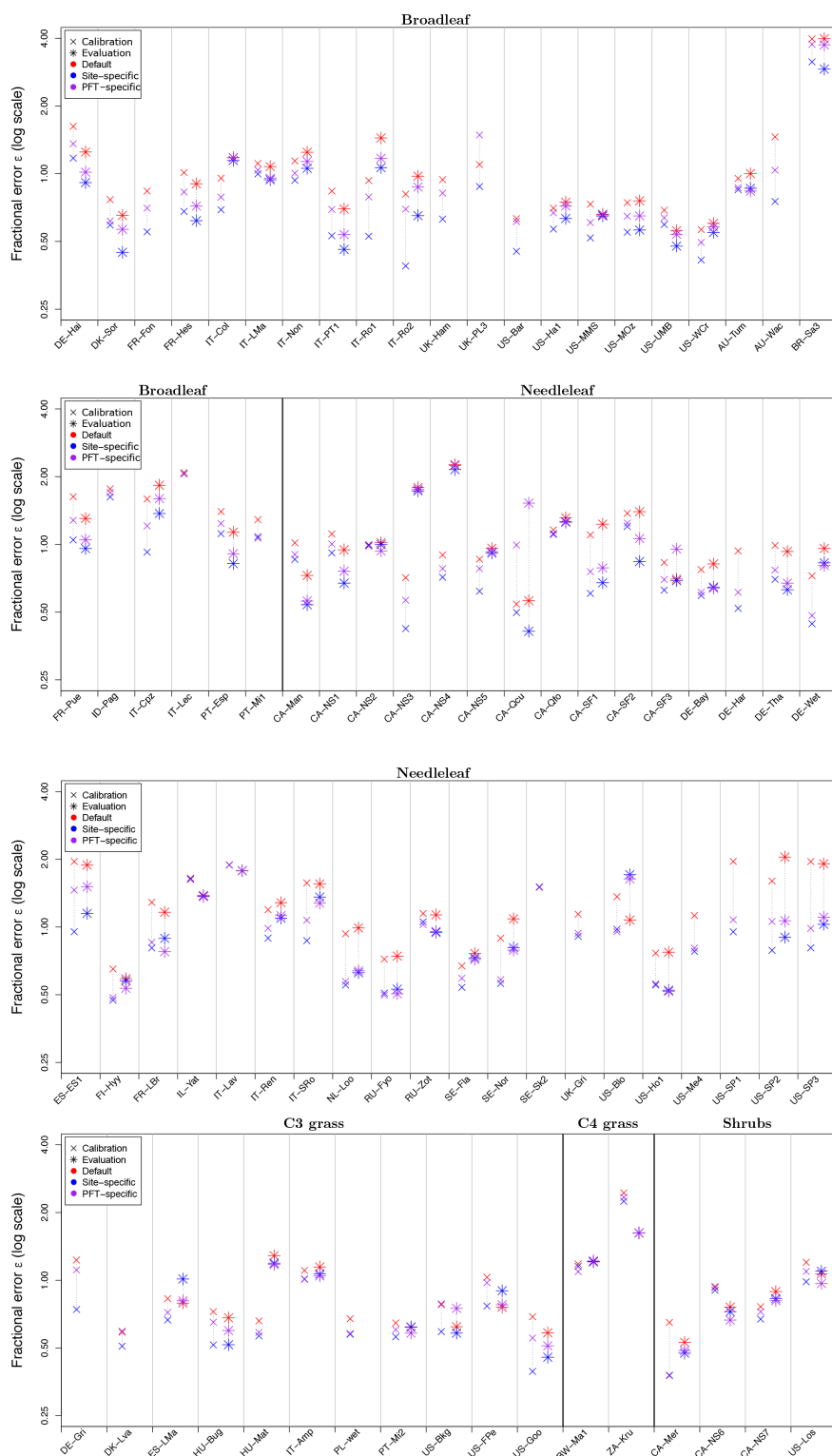


**Figure 4.** The correlations between parameters for PFT-specific parameter optimisations. Each subfigure shows a 2-D correlation map, within which each box is a 2-D marginal plot. Bar graphs show 1-D marginal distributions for individual parameters. The dimensions of the boxes represent the prior range of each parameter. Red points/dashed lines represent initial parameter values. Blue points/dashed lines represent optimised parameter values. Blue contours illustrate the posterior distribution.

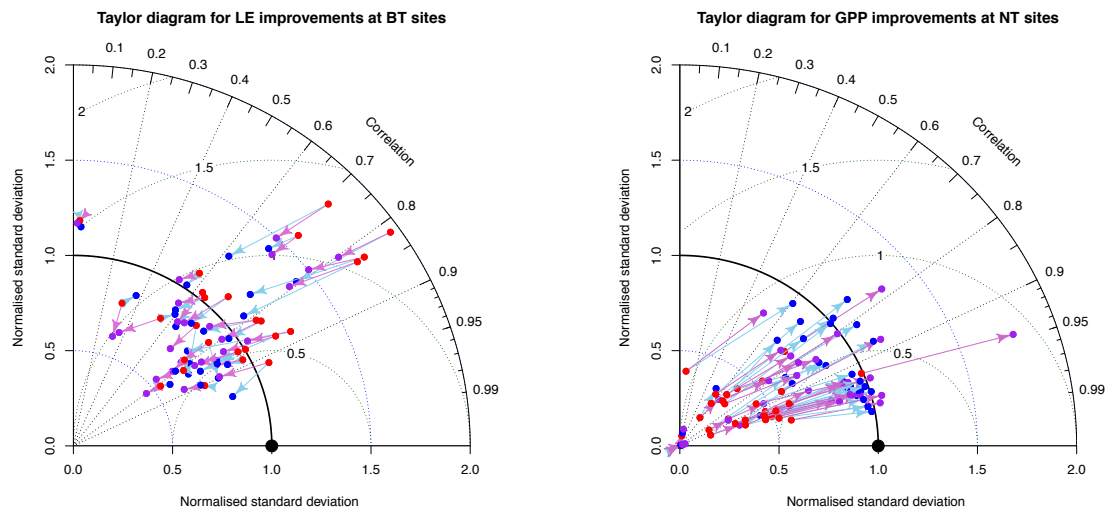
sites, only UK-PL3 gets notably worse. Investigation shows that this site behaves differently from the rest of the sites in the set, both in the magnitude of the fluxes and seasonality. This UK site is in the Pang–Lambourn catchment, which has chalk soil with macropores that permit significant lateral sub-surface flows of soil moisture. These horizontal flows cannot

be captured in a model like JULES, which is essentially 1-D in the vertical below the soil surface.

Similar levels of fit and error reduction can be seen in the evaluation years in the broadleaf set. Only IT-Col and US-MMS show no improvement, the PFT-specific optimal parameter vector does not worsen the fit at these locations.



**Figure 5.** Calibration and evaluation of site-specific and PFT-specific parameter optimisation at FLUXNET sites, using the metric described in Sect. 2.5.2. Fractional error shown for default JULES parameters (red), site-specific optimal parameters (blue), PFT-specific optimal parameters (violet). Results are shown both for the calibration year ( $\times$ , on left) and for the evaluation year ( $*$ , on right). No evaluation year was available for some sites (broadleaf: FR-Fon, UK-Ham, UK-PL3, US-Bar, ID-Pag, IT-Lec, PT-Mi1; needleleaf: SE-Sk2, UK-Gri, US-Me4, US-SP1; shrubs: DE-Gri, DK-Lva, PL-wet). Sites with very large initial errors have been removed from the plot (broadleaf: BR-Sal1; shrubs: IT-Pia).



**Figure 6.** Improvements in fit represented by “Taylor diagrams”. Observed time series (black dot) can be compared with modelled time series for default parameters (red dots), site-specific optimal parameters (blue dots) and PFT-generic optimal parameters (purple dots). Radial distance from the origin (dotted lines) represents normalised standard deviation  $\sqrt{\text{var}(\mathbf{m}_t)/\text{var}(\mathbf{o}_t)}$ , and so a modelled time series with the correct variance lies on the thick black line. Angular position represents the correlation between modelled and observed time series. The distance from the black dot (dotted green lines) represents the normalised standard deviation in the errors  $\sqrt{\text{var}(\mathbf{o}_t - \mathbf{m}_t)/\text{var}(\mathbf{o}_t)}$ .

For AU-Tum, the PFT-specific parameter vector outperforms the site-specific vector. This illustrates that the PFT-specific vector can be robust, whereas the locally optimised vectors might over-tune to the specific behaviour of the calibration year.

Results are similar for the needleleaf sites, the majority of the sites show noticeable improvements in both the calibration and evaluation years when using site-specific optimal parameter vectors. For over a third of the sites in this PFT, the improvement when using the PFT-specific parameter vector is similar to that obtained with the site-specific parameter vector. This illustrates that these sites fit well together as a single PFT. For these sites, the PFT-specific vector sometimes outperforms the site-specific vector on the evaluation years. Some sites in the needleleaf PFT remain unchanged regardless of the parameter vector used. Anomalous sites that should be noted are CA-Qcu, CA-SF3 and US-Blo. The CA-Qcu site is the only one in this PFT that does not improve when using the PFT-specific vector, for either the calibration or evaluation years. This site has a lower annual cycle of GPP than the rest in this set. The CA-SF3 site improves when using the site-specific parameter vector in the evaluation year, but not using the PFT-specific vector. The US-Blo site improves in the calibration year, but when confronted with the evaluation year, both the site-specific vector and PFT-specific vector worsen the fit. This evaluation year has unusually high LE, which might be causing this discrepancy.

For some sites, (e.g. US-Blo and BW-Ma1), the PFT-specific optimum outperforms the site-specific optimum in the calibration year. This phenomenon was also noted by Kuppel et al. (2014), who suggest that the added constraints

placed on the parameters by increasing the number of sites causes the cost function to become “smoother”. This may render the optimisation scheme less likely to become trapped in local minima.

The last panel of Fig. 5 shows the C3 grass sites, the C4 grass sites and the shrubs sites. For the C3 grass sites, the majority of the evaluation years have a better fit with the PFT-specific parameter vector than with site-specific parameter vector. This suggests that the seasonal cycle differs over the different years at these sites. For the C4 grass sites, which started with relatively high errors, the new parameter vectors improve the sites slightly for the calibration year but hardly at all for the evaluation year. This set of two sites is too small to draw any proper conclusion about the C4 grass parameters. There is a clear need for more data from C4 grass sites. Finally, the shrubs can be seen to improve for all the sites. For the shrub sites, both the site-specific and the PFT-specific provide a better fit of the model to the observations of the calibration year. The improvement is minor for these sites, except for CA-Mer, which halves its fractional error. When confronted with observations from the evaluation years, the model also improves the fit of these sites for both site-specific and PFT-specific parameters (with the exception of US-Los, where the site-specific optimal vector increases error but the PFT-specific vector reduces it). This is another example of the PFT-specific parameter vector being more robust.

### 3.4 Analysis of improvement in fit

The fractional error is a good tool for cross-site comparison but it does not give much information about the way in which the optimised parameter vectors improve the fit at each site.

Taylor diagrams (Taylor, 2001) provide more insight into how the fit has been improved by considering the relationship between observed variance  $\text{var}(o_t)$ , modelled variance  $\text{var}(m_t)$ , error variance  $\text{var}(o_t - m_t)$  and model–observation correlation  $\text{cor}(o_t, m_t)$ .

The Taylor diagrams in Fig. 6 illustrate the improvement in performance of the optimised model for both the site-specific and PFT-generic parameters during calibration years (plots for evaluation years are very similar).

For latent heat at broadleaf sites (left), the improvement is most noticeable in cases where the seasonal cycle was over-estimated. The correlation between the modelled time series and observation time series does not improve much but for the majority of the sites this starts off relatively high (over 0.6). Other PFTs show less drastic improvements for latent heat.

For GPP at needleleaf sites (right), the seasonal cycle is typically underestimated and improves noticeably for both the single-site parameter vectors and the PFT-generic parameter vectors. The correlation between model and observed time series does not change greatly. The Taylor diagram for GPP at broadleaf sites is very similar. For grasses and shrubs, the change is less drastic, though some of the sites have a more notable increase in correlation.

Since Taylor diagrams are based on a decomposition of the variance of the errors they are insensitive to any systematic offset in the model. It therefore makes sense to consider in addition the normalised bias  $|\mu_m - \mu_o|/\sigma_o$ . Calculating this statistic separately shows a reduction in bias for nearly all sites. Taken together, these measures show that the observed improvements in model fit are due mainly to adjustment of the magnitude of the annual cycle and a reduction in bias.

## 4 Conclusions

This study introduces the adJULES system, which has been developed to tune the internal parameters of the JULES land-surface model. adJULES enables objective calibration of JULES against observational data, providing best-fit internal parameters and associated uncertainty ranges.

For individual FLUXNET sites, adJULES has the ability to find local (site-specific) optimal parameter vectors that significantly improve the performance of the JULES model compared to runs generated using the default parameters. The data streams used in the calibration, LE and GPP, are both modelled more accurately with the optimal parameter vectors, with the GPP flux improving the most. The greater improvement in the GPP flux is due to the fact that the parameters considered in this study are mainly related to photosynthesis. For the LE flux to improve more significantly, more water and energy-related parameters would need to be considered in the optimisation.

When optimised locally to find site-specific parameters, all of the sites in this study improve the model–data fit for

the calibration year. In addition, when confronted with independent data from an evaluation year, the locally optimised parameter vectors decreased the error in model–data fit for 92 % of the sites. This evaluation of the site-specific parameter vectors is promising, and suggests that the adJULES system is robust. It also gives confidence that the parameter vectors found can be generalised over different locations.

This study is motivated partly by the desire to improve the performance of JULES within the Hadley Centre's Earth system models, which means needing to find best-fit parameters for a relatively small number of PFTs. The adJULES system has the ability to calibrate multiple locations simultaneously in order to find best-fit parameters. This multi-site optimisation is a relatively new feature in terrestrial data assimilation. By classifying the FLUXNET sites into groups dominated by each JULES PFT (BT, NT, C3G, C4G, Sh), adJULES was used to find the optimal PFT-specific parameters.

Although the PFT-specific optimal parameters do not always fit the data as well as site-specific optimal parameters, they still offer significant improvements over the default JULES parameters. For over 85 % of the sites, PFT-specific optimal parameters perform better than default parameters when confronted with independent evaluation data. For 50 % of the sites, the PFT-specific optimal parameters perform at least as well as site-specific optimal parameters. This implies that the multi-site methodology is less susceptible to over-tuning, both in terms of variability across sites (e.g. different overground biomass and tree ranges), and in terms of variability through time (e.g. unusually high rainfall in the calibration year).

The PFT-specific parameters found in this study represent a significant improvement on the default ones. The fact that such parameters could be found implies robust parameterisations independent of geography. This supports the idea that it is possible to represent global vegetation with a relatively small number of PFTs.

A successful and robust multi-site optimisation assumes that sites can be grouped and parameter values can apply to several sites at once. Whilst the PFT-specific parameters show great improvement, agreeing with the use of five PFTs in JULES, it would be possible to rethink the PFT definitions and group sites differently. This could be done either by looking more closely at the site specifics detailed in the FLUXNET database, or by considering single-site optimisations and performing a cluster analysis in parameter space to identify PFTs empirically.

There is always a risk of becoming stuck in local minima when optimising within a high-dimensional parameter space by gradient descent. When an optimisation finds a local minimum, the final optimised state depends on the initial conditions. The consistency between our single site and multiple site optimisations therefore gives us some confidence in the robustness of the convergence of our algorithm for this application.

It is however clear that there are some limitations to the success of the optimisation results. Some sites still show significant differences between model output and observations. This suggests that an improvement to model physics may be necessary in order to produce better model output. This is because adJULES produces the (locally) best possible fit to observations, given the existing model physics and the prescribed driving data. If the fit is still inadequate, this may be due to the model and data themselves, rather than parameter values. adJULES can therefore be used in the identification of model structural errors. Another reason for inadequate fit may be due to the method used. A limitation of gradient descent methods, such as the optimisation scheme used in this study, is that the local minimum found depends on the initial parameter vector. However, as discussed in Sect. 3.3, the fact that the cost function becomes smoother with additional sites may help with becoming trapped in local minima (Kuppel et al., 2014). Alternative methods, including ensemble methods, could avoid this issue, but are more computationally costly. For some PFTs (notably C4G and shrubs) there are insufficient FLUXNET sites to determine optimal parameters satisfactorily. Additional data and sites for these PFTs are therefore urgently required.

## 5 Code availability

The source code of the adJULES data assimilation system is available at <http://adjules.ex.ac.uk/>. The JULES land-surface model is freely available to any researcher for non-commercial use. Version 2.2 used in this study can be requested at <http://jules.jchmr.org>. The main documentation for the JULES system can also be found at this site. The adjoint of the JULES model has been generated using commercial software TAF (Sect. 2.2.1). For licensing reasons, the recalculation of the adjoint following code changes can be done only by the authors at the University of Exeter.

## Appendix A

**Table A1.** FLUXNET sites used in this study, labelled by a country code (first two letters) and site name (last three letters). The period corresponds to the available years of data for each of the sites.

Site	Period	Calibration year	Evaluation year	Latitude	Longitude
Broadleaf sites (BT)					
DE-Hai	(2000, 2006)	2005	2004	51.079	10.452
DK-Sor	(1996, 2006)	2006	2004	55.487	11.646
FR-Fon	(2005, 2006)	2006	–	48.476	2.780
FR-Hes	(1997, 2006)	2003	1998	48.674	7.065
IT-Col	(1996, 2006)	2005	2001	41.849	13.588
IT-LMa	(2003, 2006)	2006	2004	45.581	7.155
IT-Non	(2001, 2006)	2002	2003	44.690	11.089
IT-PT1	(2002, 2004)	2003	2004	45.201	9.061
IT-Ro1	(2000, 2006)	2006	2005	42.408	11.930
IT-Ro2	(2002, 2006)	2004	2006	42.390	11.921
UK-Ham	(2004, 2005)	2005	–	51.121	–0.861
UK-PL3	(2005, 2006)	2006	–	51.450	–1.267
US-Bar	(2004, 2005)	2005	–	44.065	–71.288
US-Ha1	(1991, 2006)	1996	1998	42.538	–72.171
US-MMS	(1999, 2005)	2002	2003	39.323	–86.413
US-MOz	(2004, 2006)	2006	2005	38.744	–92.200
US-UMB	(1999, 2003)	2003	2002	45.560	–84.714
US-WCr	(1999, 2006)	2005	2000	45.806	–90.080
AU-Tum	(2001, 2006)	2003	2005	–35.656	148.152
AU-Wac	(2005, 2007)	2006	–	–37.429	145.187
BR-Sa1	(2002, 2004)	2003	2004	–2.857	–54.959
BR-Sa3	(2000, 2003)	2002	2003	–3.018	–54.971
FR-Pue	(2000, 2006)	2006	2005	43.741	3.596
ID-Pag	(2002, 2003)	2003	–	2.345	114.036
IT-Cpz	(1997, 2006)	2004	2006	41.705	12.376
IT-Lec	(2005, 2006)	2006	–	43.305	11.271
PT-Esp	(2002, 2004)	2004	2003	38.639	–8.602
PT-Mi1	(2003, 2005)	2005	–	38.541	–8.000
C3 grasses sites (C3G)					
DE-Gri	(2005, 2006)	2006	–	50.950	13.512
DK-Lva	(2005, 2006)	2006	–	55.683	12.083
ES-LMa	(2004, 2006)	2006	2005	39.941	–5.773
HU-Bug	(2002, 2006)	2006	2005	46.691	19.601
HU-Mat	(2004, 2006)	2006	2005	47.847	19.726
IT-Amp	(2002, 2006)	2006	2005	41.904	13.605
PL-wet	(2004, 2005)	2005	–	52.762	16.309
PT-Mi2	(2004, 2006)	2006	2005	38.477	–8.025
US-Bkg	(2004, 2006)	2006	2005	44.345	–96.836
US-FPe	(2000, 2006)	2002	2004	48.308	–105.101
US-Goo	(2002, 2006)	2006	2004	34.250	–89.970

Table A1. Continued.

Site	Period	Calibration year	Evaluation year	Latitude	Longitude
Needleleaf sites (NT)					
CA-Man	(1997, 2003)	2001	2002	55.880	−98.481
CA-NS1	(2002, 2005)	2004	2003	55.879	−98.484
CA-NS2	(2001, 2005)	2002	2004	55.906	−98.525
CA-NS3	(2001, 2005)	2004	2002	55.912	−98.382
CA-NS4	(2002, 2004)	2004	2003	55.912	−98.382
CA-NS5	(2001, 2005)	2004	2002	55.863	−98.485
CA-Qcu	(2001, 2006)	2005	2006	49.267	−74.037
CA-Qfo	(2003, 2006)	2006	2005	49.693	−74.342
CA-SF1	(2003, 2005)	2004	2005	54.485	−105.818
CA-SF2	(2003, 2005)	2004	2005	54.254	−105.878
CA-SF3	(2003, 2005)	2005	2004	54.092	−106.005
DE-Bay	(1996,1999)	1999	1998	50.142	11.867
DE-Har	(2005, 2006)	2006	–	47.934	7.601
DE-Tha	(1996, 2006)	2005	2004	50.964	13.567
DE-Wet	(2002, 2006)	2006	2004	50.453	11.457
ES-ES1	(1999, 2006)	2005	2000	39.346	−0.319
FI-Hyy	(1996, 2006)	2006	2004	61.847	24.295
FR-LBr	(2003, 2006)	2006	2005	44.717	−0.769
IL-Yat	(2001, 2006)	2005	2006	31.345	35.051
IT-Lav	(2000, 2002)	2001	2002	45.955	11.281
IT-Ren	(1999, 2006)	2005	2006	46.588	11.435
IT-SRo	(1999, 2006)	2006	2005	43.728	10.284
NL-Loo	(1996, 2006)	2006	2003	52.168	5.744
RU-Fyo	(1998, 2006)	2005	2006	56.462	32.924
RU-Zot	(2002, 2004)	2003	2004	60.801	89.351
SE-Fla	(1996,1998)	1998	1997	64.113	19.457
SE-Nor	(1996,1999)	1997	1999	60.086	17.480
SE-Sk2	(2004, 2005)	2005	–	60.130	17.840
UK-Gri	(1997,1998)	1998	–	56.607	−3.798
US-Blo	(1997, 2006)	2006	2000	38.895	−120.633
US-Ho1	(1996, 2004)	2004	2003	45.204	−68.740
US-Me4	(1996, 2000)	2000	–	44.499	−121.622
US-SP1	(2000, 2001)	2001	–	29.738	−82.219
US-SP2	(1998, 2004)	2001	2004	29.765	−82.245
US-SP3	(1999, 2004)	2001	2002	29.755	−82.163
Shrubs sites (Sh)					
CA-Mer	(1998, 2005)	2004	2005	45.409	−75.519
CA-NS6	(2001, 2005)	2003	2004	55.917	−98.964
CA-NS7	(2002, 2005)	2003	2004	56.636	−99.948
IT-Pia	(2002, 2005)	2003	2004	42.584	10.078
US-Los	(2001, 2005)	2005	2003	46.083	−89.979
C4 grasses sites (C4G)					
BW-Ma1	(1999, 2001)	2000	2001	−19.916	23.561
ZA-Kru	(2001, 2003)	2002	2003	−25.020	31.497



Appendix B

**Table B1.** PFT-specific JULES parameters optimised in this study (Table 1). The prior values and ranges for each PFT are given. Below in bold are the optimised values and posterior uncertainty ranges given as an 80 % confidence interval (in parentheses). Optimised values for which the prior values lie outside the posterior range are highlighted by (\*). A graphical version of this table is shown in Fig. 3.

	BT	NT	C3	C4	Sh
$n_0$	0.046 (0.001, 0.2) <b>0.061</b> <b>(0.034, 0.066)</b>	0.033 (0.001, 0.2) <b>0.065*</b> <b>(0.059, 0.07)</b>	0.073 (0.001, 0.2) <b>0.07</b> <b>(0.018, 0.145)</b>	0.06 (0.001, 0.2) <b>0.051*</b> <b>(0.043, 0.056)</b>	0.06 (0.001, 0.2) <b>0.041</b> <b>(0.006, 0.066)</b>
$\alpha$	0.08 (0.001, 0.999) <b>0.131*</b> <b>(0.087, 0.14)</b>	0.08 (0.001, 0.999) <b>0.096</b> <b>(0.021, 0.167)</b>	0.12 (0.001, 0.999) <b>0.179*</b> <b>(0.155, 0.209)</b>	0.06 (0.001, 0.999) <b>0.118*</b> <b>(0.075, 0.141)</b>	0.08 (0.001, 0.999) <b>0.102</b> <b>(0.063, 0.763)</b>
$f_0$	0.875 (0.5, 0.99) <b>0.765*</b> <b>(0.655, 0.787)</b>	0.875 (0.5, 0.99) <b>0.737*</b> <b>(0.713, 0.758)</b>	0.9 (0.5, 0.99) <b>0.817</b> <b>(0.727, 0.944)</b>	0.8 (0.5, 0.99) <b>0.765*</b> <b>(0.752, 0.793)</b>	0.9 (0.5, 0.99) <b>0.782*</b> <b>(0.735, 0.848)</b>
$T_{low}$	0 (-50, 40) <b>1.203</b> <b>(-0.555, 9.492)</b>	-10 (-50, 40) <b>-8.698</b> <b>(-10.98, -6.342)</b>	0 (-50, 40) <b>-1.985*</b> <b>(-3.877, -0.13)</b>	13 (-50, 40) <b>11.37</b> <b>(7.522, 14.072)</b>	0 (-50, 40) <b>-5.208*</b> <b>(-10.855, -2.106)</b>
$T_{upp}$	36 (25, 50) <b>38.578*</b> <b>(38.157, 40.698)</b>	26 (25, 50) <b>34.721*</b> <b>(33.214, 36.365)</b>	36 (25, 50) <b>36.242</b> <b>(33.087, 38.599)</b>	45 (25, 50) <b>44.897</b> <b>(44.201, 46.426)</b>	36 (25, 50) <b>35.385</b> <b>(26.339, 40.216)</b>
$d_r$	3 (0.1, 4) <b>3.009</b> <b>(2.901, 3.052)</b>	1 (0.1, 4) <b>1.425*</b> <b>(1.159, 1.672)</b>	0.5 (0.1, 4) <b>0.991*</b> <b>(0.901, 1.101)</b>	0.5 (0.1, 4) <b>0.404*</b> <b>(0.5, 3.623)</b>	0.5 (0.1, 4) <b>0.411*</b> <b>(0.324, 0.473)</b>
$\frac{\delta c}{\delta t}$	0.05 (0.001, 0.1) <b>0.047*</b> <b>(0.046, 0.049)</b>	0.05 (0.001, 0.1) <b>0.045*</b> <b>(0.042, 0.048)</b>	0.05 (0.001, 0.1) <b>0.05</b> <b>(0.047, 0.052)</b>	0.05 (0.001, 0.1) <b>0.05</b> <b>(0.046, 0.054)</b>	0.05 (0.001, 0.1) <b>0.048</b> <b>(0.04, 0.055)</b>
$dq_c$	0.09 (0.001, 0.2) <b>0.048</b> <b>(0.02, 0.183)</b>	0.06 (0.001, 0.2) <b>0.036</b> <b>(0.008, 0.066)</b>	0.1 (0.001, 0.2) <b>0.086</b> <b>(0.07, 0.109)</b>	0.075 (0.001, 0.2) <b>0.046*</b> <b>(0.045, 0.053)</b>	0.1 (0.001, 0.2) <b>0.077</b> <b>(0.024, 0.118)</b>

**Acknowledgements.** This work was supported by the UK Natural Environment Research Council (NERC) through the National Centre for Earth Observation (NCEO).

This study used eddy-covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCAN), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. Support for eddy-covariance data harmonisation was provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California – Berkeley, University of Virginia.

The authors are grateful to T. Kaminski and R. Giering from FastOpt for their contribution to the development of the adjoint model, and to M. Groenendijk, A. Harper, and the UK Met Office for processing and sharing their data.

The authors are particularly grateful to two anonymous referees for their thoughtful and constructive reviews, which greatly improved this manuscript.

Edited by: J. Kala

Reviewed by: H. Gupta and three anonymous referees

## References

- Ajami, N. K., Duan, Q., and Sorooshian, S.: An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res.*, 43, w01403, doi:10.1029/2005WR004745, 2007.
- Arora, V. and Boer, G.: A parameterization of leaf phenology for the terrestrial ecosystem component of climate models, *Glob. Change Biol.*, 11, 39–59, 2005.
- Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *B. Am. Meteorol. Soc.*, 82, 2415–2434, 2001.
- Bartholomew-Biggs, M., Brown, S., Christianson, B., and Dixon, L.: Automatic differentiation of algorithms, *J. Comput. Appl. Math.*, 124, 171–190, doi:10.1016/S0377-0427(00)00422-2, 2000.
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, doi:10.5194/gmd-4-677-2011, 2011.
- Blyth, E., Lloyd, A., Gash, J., Pryor, M., Weedon, G., and Shuttleworth, J.: Evaluating the JULES Land Surface Model Energy Fluxes Using FLUXNET Data, *J. Hydrometeorol.*, 11, 509–519, 2010.
- Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., and Bellouin, N.: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability, *Nature*, 484, 228–232, 2012.
- Bouttier, F. and Courtier, P.: Data assimilation concepts and methods, Training, 1–59, Meteorological training course lecture series, ECMWF, 1999.
- Braswell, B. H., Sacks, W. J., Linder, E., and Schimel, D. S.: Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, *Glob. Change Biol.*, 11, 335–355, doi:10.1111/j.1365-2486.2005.00897.x, 2005.
- Brovkin, V., Boysen, L., Raddatz, T., Gayler, V., Loew, A., and Claussen, M.: Evaluation of vegetation cover and land-surface albedo in MPI-ESM CMIP5 simulations, *J. Adv. Model. Earth Syst.*, 5, 48–57, doi:10.1029/2012MS000169, 2013.
- Byrd, R., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained optimization, *SIAM Journal on Scientific Computing*, 16, 1190–1208, 1995.
- Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, *Geosci. Model Dev.*, 4, 701–722, doi:10.5194/gmd-4-701-2011, 2011.
- Collatz, G., Ball, J., Grivet, C., and Berry, J.: Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer, *Agr. Forest Meteorol.*, 54, 107–136, 1991.
- Collatz, G. J., Ribas-Carbo, M., and Berry, J.: Coupled photosynthesis-stomatal conductance model for leaves of C4 plants, *Funct. Plant Biol.*, 19, 519–538, 1992.
- Cox, P., Huntingford, C., and Harding, R.: A canopy conductance and photosynthesis model for use in a GCM land surface scheme, *J. Hydrol.*, 212, 79–94, 1998.
- Cox, P., Betts, R., Jones, C., Spall, S., and Totterdell, I.: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, 408, 184–187, 2000.
- Cox, P. M., Betts, R. A., Bunton, C. B., Essery, R. L. H., Rowntree, P. R., and Smith, J.: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, *Clim. Dynam.*, 15, 183–203, 1999.
- Friedlingstein, P., Bopp, L., Ciais, P., Dufresne, J.-L., Fairhead, L., LeTreut, H., Monfray, P., and Orr, J.: Positive feedback between future climate change and the carbon cycle, *Geophys. Res. Lett.*, 28, 1543–1546, doi:10.1029/2000GL012015, 2001.
- Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, F., Fung, I. M., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H., Radiate, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A., Yoshikawa, C., and Zeng, N.: Climate–carbon cycle

- feedback analysis: results from the C4MIP model intercomparison, *J. Climate*, 19, 3337–3353, 2006.
- Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, *J. Climate*, 27, 511–526, doi:10.1175/JCLI-D-12-00579.1, 2013.
- Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE T. Pattern Anal.*, 6, 721–741, 1984.
- Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., and Hothorn, T.: mvtnorm: Multivariate Normal and t Distributions, R package version 1.0-5, available at: <http://CRAN.R-project.org/package=mvtnorm> (last access: 23 August 2016), 2015.
- Giering, R., Kaminski, T., and Slawig, T.: Generating efficient derivative code with TAF: Adjoint and tangent linear Euler flow around an airfoil, *Future Gener. Comp. Sy.*, 21, 1345–1355, doi:10.1016/j.future.2004.11.003, 2005.
- Groenendijk, M., Dolman, A., van der Molen, M., R. Luning, A. A., Delpierre, N., Gash, J., Lindorth, A., Richardson, A., Verbeeck, H., and Wohlfahrt, G.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global FLUXNET eddy covariance data, *Agr. Forest Meteorol.*, 151, 22–38, 2010.
- Gupta, H. V., Bastidas, L. A., Sorooshian, S., Shuttleworth, W. J., and Yang, Z. L.: Parameter estimation of a land surface scheme using multicriteria methods, *J. Geophys. Res.-Atmos.*, 104, 19491–19503, doi:10.1029/1999JD900154, 1999.
- Hararuk, O., Xia, J., and Luo, Y.: Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method, *J. Geophys. Res.-Biogeo.*, 119, 403–417, doi:10.1002/2013JG002535, 2014.
- Jones, G. S., Stott, P. A., and Christidis, N.: Attribution of observed historical near-surface temperature variations to anthropogenic and natural causes using CMIP5 simulations, *J. Geophys. Res.-Atmos.*, 118, 4001–4024, doi:10.1002/jgrd.50239, 2013.
- Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., van Scheltinga, A. T., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, *J. Geophys. Res.-Biogeo.*, 118, 1414–1426, doi:10.1002/jgrg.20118, 2013.
- Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, 42, w03407, doi:10.1029/2005WR004368, 2006a.
- Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resour. Res.*, 42, w03408, doi:10.1029/2005WR004376, 2006b.
- Knorr, W. and Kattge, J.: Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, *Glob. Change Biol.*, 11, 1333–1351, 2005.
- Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, *Biogeosciences*, 9, 3757–3776, doi:10.5194/bg-9-3757-2012, 2012.
- Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., and Cescatti, A.: Model–data fusion across ecosystems: from multisite optimizations to global simulations, *Geosci. Model Dev.*, 7, 2581–2597, doi:10.5194/gmd-7-2581-2014, 2014.
- Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, 43, w07401, doi:10.1029/2006WR005756, 2007.
- Madsen, H.: Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Adv. Water Resour.*, 26, 205–216, doi:10.1016/S0309-1708(02)00092-1, 2003.
- Medvigy, D. and Moorcroft, P. R.: Predicting ecosystem dynamics at regional scales: an evaluation of a terrestrial biosphere model for the forests of northeastern North America, *Philos. T. R. Soc. Lond. B*, 367, 222–235, doi:10.1098/rstb.2011.0253, 2011.
- Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y., and Moorcroft, P. R.: Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2, *J. Geophys. Res.-Biogeo.*, 114, g01002, doi:10.1029/2008JG000812, 2009.
- Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., and Cox, P. M.: Impact of changes in diffuse radiation on the global land carbon sink, *Nature*, 458, 1014–1017, 2009.
- Myneni, R., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G., Lotsch, A., Friedl, M., Morisette, J., Votava, P., Nemaini, R., and Running, S.: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sens. Environ.*, 83, 214–231, 2002.
- Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.
- Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, *Biogeosciences*, 3, 571–583, doi:10.5194/bg-3-571-2006, 2006.
- Peng, C., Guilot, J., Wu, H., Jiang, H., and Luo, Y.: Integrating models with data in ecology and palaeoecology advances towards a model-data fusion approach, *Ecol. Lett.*, 14, 522–536, 2011.
- Pitman, A.: The evolution of, and revolution in, land surface schemes designed for climate models, *Int. J. Climatol.*, 23, 479–510, 2003.
- Raupach, M., Rayner, P., Barrett, D., Defries, R., Heimann, M., Ojima, D., Quegan, S., and Schmullius, C.: Model-data synthesis in terrestrial carbon observation methods, data requirements and data uncertainty specifications, *Glob. Change Biol.*, 11, 378–297, 2005.
- Rayner, P., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), *Glob. Biogeochem. Cy.*, 19, GB2026, doi:10.1029/2004GB002254, 2005.

- R Development Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, 2015.
- Reichstein, M., Tenhunen, J., Rouspard, O., Ourcival, J.-M., Rambal, S., Miglietta, F., Peressotti, A., Pecchiari, M., Tirone, G., and Valentini, R.: Inverse modeling of seasonal drought effects on canopy CO<sub>2</sub>/H<sub>2</sub>O exchange in three Mediterranean ecosystems, *J. Geophys. Res.-Atmos.*, 108, 4726, doi:10.1029/2003JD003430, 2003.
- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, 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., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, *Glob. Change Biol.*, 11, 1424–1439, 2005.
- Santaren, D., Peylin, P., Viovy, N., and Ciais, P.: Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France, *Global Biogeochem. Cy.*, 21, GB2013, doi:10.1029/2006GB002834, 2007.
- Sellers, P., Dickinson, R., Randall, D., Betts, A., Hall, F., Berry, J., Collatz, G., Denning, A., Mooney, H., Nobre, C., and Sato, N.: Modeling the exchanges of energy, water, and carbon between continents and the atmosphere, *Science*, 275, 502–509, 1997.
- Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. (Eds.): Summary for Policymakers, book section SPM, 1–30, Cambridge University Press, Cambridge, UK and New York, NY, USA, doi:10.1017/CBO9781107415324.004, 2013.
- Stöckli, R., Rutishauser, T., Dragoni, D., O’Keefe, J., Thornton, P. E., Jolly, M., Lu, L., and Denning, A. S.: Remote sensing data assimilation for a prognostic phenology model, *J. Geophys. Res.-Biogeo.*, 113, g04021, doi:10.1029/2008JG000781, 2008.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.-Atmos.*, 106, 7183–7192, doi:10.1029/2000JD900719, 2001.
- Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO<sub>2</sub> fertilization and climate variability, *Global Biogeochem. Cy.*, 21, gB4018, doi:10.1029/2006GB002868, 2007.
- Thum, T., Aalto, T., Laurila, T., Aurela, M., Lindroth, A., and Vesala, T.: Assessing seasonality of biochemical CO<sub>2</sub> exchange model parameters from micrometeorological flux observations at boreal coniferous forest, *Biogeosciences*, 5, 1625–1639, 2008.
- Verbeeck, H., Peylin, P., Bacour, C., Bonal, D., Steppe, K., and Ciais, P.: Seasonal patterns of CO<sub>2</sub> fluxes in Amazon forests: Fusion of eddy covariance data and the ORCHIDEE model, *J. Geophys. Res.*, 116, G02018, doi:10.1029/2010JG001544, 2011.
- Wang, Y.-P., Leuning, R., Cleugh, H. A., and Coppin, P. A.: Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful?, *Glob. Change Biol.*, 7, 495–510, 2001.
- Wang, Y. P., Baldocchi, D., Leuning, R., Falge, E., and Vesala, T.: Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight Fluxnet sites, *Glob. Change Biol.*, 13, 652–670, 2007.
- Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y.-P.: Improving land surface models with FLUXNET data, *Biogeosciences*, 6, 1341–1359, doi:10.5194/bg-6-1341-2009, 2009.
- Xiao, J., Davis, K. J., Urban, N. M., Keller, K., and Saliendra, N. Z.: Upscaling carbon fluxes from towers to the regional scale: Influence of parameter variability and land cover representation on regional flux estimates, *J. Geophys. Res.-Biogeo.*, 116, G00J06, doi:10.1029/2010JG001568, 2011.
- Zaehle, S., Friedlingstein, P., and Friend, A. D.: Terrestrial nitrogen feedbacks may accelerate future climate change, *Geophys. Res. Lett.*, 37, 101401, doi:10.1029/2009GL041345, 2010.