



Multi-site evaluation of the JULES land surface model using global and local data

D. Slevin, S. F. B. Tett, and M. Williams

School of Geosciences, The University of Edinburgh, Grant Institute, James Hutton Road, Edinburgh, EH9 3FE, UK

Correspondence to: D. Slevin (d.slevin@ed.ac.uk)

Received: 8 July 2014 – Published in Geosci. Model Dev. Discuss.: 8 August 2014

Revised: 15 January 2015 – Accepted: 20 January 2015 – Published: 13 February 2015

Abstract. This study evaluates the ability of the JULES land surface model (LSM) to simulate photosynthesis using local and global data sets at 12 FLUXNET sites. Model parameters include site-specific (local) values for each flux tower site and the default parameters used in the Hadley Centre Global Environmental Model (HadGEM) climate model. Firstly, gross primary productivity (GPP) estimates from driving JULES with data derived from local site measurements were compared to observations from the FLUXNET network. When using local data, the model is biased with total annual GPP underestimated by 16 % across all sites compared to observations. Secondly, GPP estimates from driving JULES with data derived from global parameter and atmospheric reanalysis (on scales of 100 km or so) were compared to FLUXNET observations. It was found that model performance decreases further, with total annual GPP underestimated by 30 % across all sites compared to observations. When JULES was driven using local parameters and global meteorological data, it was shown that global data could be used in place of FLUXNET data with a 7 % reduction in total annual simulated GPP. Thirdly, the global meteorological data sets, WFDEI and PRINCETON, were compared to local data to find that the WFDEI data set more closely matches the local meteorological measurements (FLUXNET). Finally, the JULES phenology model was tested by comparing results from simulations using the default phenology model to those forced with the remote sensing product MODIS leaf area index (LAI). Forcing the model with daily satellite LAI results in only small improvements in predicted GPP at a small number of sites, compared to using the default phenology model.

1 Introduction

The atmosphere and biosphere are closely coupled and carbon is transported between the two via the carbon cycle (Cao and Woodward, 1998). Although the carbon cycle is significantly affected by global warming, much still remains to be understood about its behaviour (Schimel, 2007). Atmospheric CO₂ represents only a small amount of carbon in the Earth system with the rest tied up in various reservoirs (Ciais et al., 2013). These reservoirs can be either sources (release more carbon than they absorb) or sinks (absorb more carbon than they release). Sources can be either man-made (combustion of fossil fuels, deforestation) or natural (plant and litter decomposition, soil respiration, ocean release) and sinks include land vegetation, soils, oceans and geological reservoirs, such as deep-sea carbonate sediments and the upper mantle (Ciais et al., 2013). Of the carbon dioxide emitted into the atmosphere from the burning of fossil fuels, roughly half remains in the atmosphere and the rest is absorbed by carbon sinks on land and in the oceans (Le Quéré et al., 2009).

Global warming can affect terrestrial ecosystems in two ways. Firstly, increasing atmospheric CO₂ concentrations have led to an increase in photosynthesis (Beck et al., 2011; Fensholt et al., 2012), which has increased both carbon uptake and storage by terrestrial ecosystems (Norby et al., 2005; Leakey et al., 2009). This is known as CO₂ fertilisation. This increase in atmospheric CO₂ has led to an increase in growing season leaf area index (LAI); see Piao et al. (2006). It also reduces plant transpiration and increases plant water use efficiency through the partial closure of stomata (Warren et al., 2011). Secondly, a warmer climate can accelerate the decomposition of litter and soil organic carbon, and increase plant respiration.

Table 1. Model parameters and meteorological variables that are altered between global and local model simulations.

Data set	Variable name	Units
Model parameters	PFT fractions	Dimensionless
	Annual maximum LAI	$\text{m}^2 \text{m}^{-2}$
	Canopy height	metres
	V_{cmax} (maximum rate of Rubisco carboxylase activity)	$\mu\text{mol CO}_2 \text{m}^{-2} \text{s}^{-1}$
	Rooting depth	metres
	Soil texture fractions ^a	% of sand, silt and clay
Meteorological data	Downward short-wave radiation	W m^{-2}
	Downward long-wave radiation	W m^{-2}
	Precipitation rate ^b	$\text{kg m}^{-2} \text{s}^{-1}$
	Surface air temperature	K
	Wind speed	m s^{-1}
	Surface air pressure	Pa
	Specific humidity	kg kg^{-1}

^a The soil texture fractions (%) are used to compute the soil hydraulic and thermal characteristics.

^b At some of the flux tower sites, the precipitation variable was separated into a rainfall rate ($\text{kg m}^{-2} \text{s}^{-1}$) and snowfall rate ($\text{kg m}^{-2} \text{s}^{-1}$).

Predictions of the future uptake of atmospheric CO_2 by the terrestrial biosphere are uncertain and this uncertainty comes from whether the terrestrial biosphere will continue to be a sink or source for CO_2 . The Coupled Climate–Carbon Cycle Model Intercomparison Project (C4MIP) was the first major study to examine the coupling between climate change and the carbon cycle (Friedlingstein et al., 2006). One of its main conclusions was the reduced efficiency of the earth system, in particular the land carbon sink, to absorb increased anthropogenic CO_2 . However, the magnitude of this effect depended on the model used.

Land surface models (LSMs) are an important component of climate models and simulate the interaction between the atmosphere and terrestrial biosphere. They represent the surface energy and water balance, climate effect of snow and carbon fluxes (Pitman, 2003) and are considered the lower boundary condition for global climate models (GCMs) (Best et al., 2011). GCMs require the carbon, water and energy fluxes between the land surface and atmosphere to be specified. Meteorological data, vegetation and soil characteristics are provided as inputs to LSMs, and using these, LSMs can predict fluxes, such as latent and sensible heat, upward long-wave radiation and net ecosystem exchange of CO_2 , which is used to determine global atmospheric CO_2 concentrations. Various LSMs have been designed over the last 40 years to calculate these fluxes (Dai et al., 2003).

The earliest GCMs to include a representation of the land surface based it on the simple “bucket” model. In this model, the soil is assumed to have a fixed water capacity (like a bucket) and at each land grid box and time step, the bucket is filled with precipitation and emptied by evaporation (Carson, 1982). The excess above its capacity is termed runoff. This model does not take vegetation or soil types into account. The

second generation of land surface schemes attempted to explicitly represent the effects of vegetation in surface energy balance calculations and include the Biosphere-Atmosphere Transfer Scheme (BATS) (three soil layers and one vegetation layer) (Dickinson, 1986) and the Simple Biosphere (SiB) Model (three soil layers and two vegetation layers) (Sellers et al., 1986). The current generation of models include the biological control of evapotranspiration with biochemical models of leaf photosynthesis linked to the biophysics of stomatal conductance (Farquhar et al., 1980; Bonan, 2008), and can respond to changes in atmospheric CO_2 in a more realistic way.

LSM components are designed using results from research literature, idealised laboratory experiments and observations from limited field campaigns (Stöckli et al., 2008; Williams et al., 2009). This can lead to sources of uncertainty in the parameterisation of processes; and as LSMs become more advanced, there is a need to understand their complexity and accuracy. LSMs can be tested in a variety of ways. Multimodel intercomparison projects provide a measure of how various LSMs behave under controlled conditions (Schaefer et al., 2012; Cadule et al., 2010; Randerson et al., 2009; Dirmeyer et al., 2006; Henderson-Sellers et al., 1996). Parameter perturbation experiments evaluate a single model and numerous simulations are performed where either one parameter is changed at a time within a given range (Knorr, 2000; Knorr and Heimann, 2001; El Maayar et al., 2002) or maximum and minimum values of parameters are used (Hallgren and Pitman, 2000). Recently, in the LSM community, there has been effort to create a more standardised form of model evaluation known as benchmarking, whereby publicly available data sets, at various temporal and spatial resolutions, along with metrics and areas of model performance to be evalu-

ated, are used by different modelling groups to test model performance (Abramowitz, 2012; Luo et al., 2012). This has previously been carried out by Abramowitz et al. (2008) and Blyth et al. (2011).

Blyth et al. (2011) evaluated JULES at 10 FLUXNET sites, representing a range of biomes and climatic conditions, where model parameter values were taken as if the model was embedded in a GCM, in order to assess the model's ability to predict observed water and carbon fluxes. We extended this work by performing model simulations whereby model parameters (Table 1) were set to observe local site conditions and were then compared to those using global and satellite data. Local site conditions were those relevant to a particular flux tower site and were obtained from the research literature, communications with site Primary Investigator and the Ameriflux data archive. Global data referred to the model parameters taken from data sets used by the global operational version of JULES and meteorological data from global gridded data sets extracted for each flux tower grid box. The satellite data referred to LAI data from the MODerate resolution Imaging Spectroradiometer (MODIS) instrument, aboard NASA's Earth Observing System (EOS) satellites, Terra and Aqua (<http://modis.gsfc.nasa.gov>).

In this study, we used 12 FLUXNET sites that cover a range of ecosystem types; temperate (6), boreal (2), mediterranean (2) and tropical (2) (Table 2), to investigate differences between using local, global and satellite-derived data sets when performing model simulations with JULES version 3.0 (Clark et al., 2011; Best et al., 2011). In particular, we wished to address the following research questions:

- How well does JULES perform when using the best available local meteorological *and* parameter data sets? Can the model simulate interannual variability?
- How well does JULES perform when using global data?
- Of the global meteorological data sets used in this study, which one compares best to FLUXNET data?
- Are improvements in simulated gross primary productivity (GPP) observed when forcing JULES with daily satellite phenology compared to using the default phenology module?

2 Methods and model

2.1 Model description

The Joint UK Land Environment Simulator (JULES) is the land surface scheme of the UK Met Office Unified Model (UM) (current version 8.6), a family of models that includes the Hadley Centre Global Environmental Model (HadGEM) climate model (<http://www.metoffice.gov.uk/research/modelling-systems/unified-model>). It has evolved from the Met Office Surface Exchange Scheme (MOSES)

(Cox et al., 1999). JULES is a mechanistic model and is able to model such processes as photosynthesis, evapotranspiration, soil and snow physics, and soil microbial activity (Blyth et al., 2011). Each model grid box is composed of nine different surface types, five of which are vegetation, referred to as plant functional types (PFTs) (broadleaf trees, needleleaf trees, C3 (temperate) grass, C4 (tropical) grass and shrubs), and four non-vegetation types (urban, inland water, bare soil and land-ice). Each grid box can be made up of the first eight surface types or is land-ice. For single-point model simulations, as used in this study, each point is treated as a grid box with data such as surface type fractions, soil texture fractions and meteorological data used as input to the model.

The surface fluxes of CO₂ associated with photosynthesis are computed on each time step for each PFT using a coupled photosynthesis-stomatal conductance model (Cox et al., 1998). These accumulated carbon fluxes are passed to TRIFFID (Top-down Representation of Interactive Foliage and Flora Including Dynamics), JULES' dynamic global vegetation model and also its terrestrial carbon cycle component (Cox, 2001). TRIFFID updates the areal coverage, LAI and canopy height for each PFT on a longer time step (usually every 10 days), based on the net carbon available to it and competition with other vegetation types (Cox, 2001). For these model simulations, vegetation competition was disabled, which meant that the PFT fractions for each site were prescribed and did not vary with time. If vegetation competition was switched on during the spin-up process, this would have introduced error into the model simulations due to unrealistic vegetation fractions.

In JULES, phenology is, typically, updated once per day by multiplying the annual maximum LAI by a scaling factor, which is calculated by using temperature-dependent leaf turnover rates (Clark et al., 2011). When calculating GPP, a multi-layer canopy was used for the scaling up of leaf-level photosynthesis to canopy level. The option used took into account the vertical gradient of canopy photosynthetic capacity (decreasing leaf nitrogen from top to bottom of canopy) and included light inhibition of leaf respiration. LAI is calculated for each canopy level (default number is 10), with a maximum LAI prescribed for each PFT. Clark et al. (2011) contains more information on the options available for the calculation of canopy photosynthesis. Two versions of JULES were used in this study. JULES3.0 is the original and publicly available release code of JULES version 3.0. The source code can be downloaded from <https://jules.jchmr.org/>. In addition, JULES3.0 was modified in order to force it with daily MODIS LAI (JULESmod). The local (stand-alone) and global operational versions of JULES are quite similar. Since UM v8.1 (using JULES v3.0), the JULES code for both have been the same with some exceptions, such as the UM/stand-alone initialisation code. The science code (e.g. photosynthesis, hydrology and soil processes) remains the same between the two. A more detailed description of JULES can be found in Clark et al. (2011) and Best et al. (2011).

Table 2. Flux towers used in this study. The following biome types were used: deciduous broadleaf forest (DBF); evergreen needleleaf forest (ENF); cropland (CRO); grassland (GRA); tundra (TUN); evergreen broadleaf forest (EBF).

Number	Site	Location		Altitude (m)	Biome type	Year	Climate zone
		Lat [° N]	Long [° E]				
1	Harvard Forest	42.54	−72.17	303	DBF	2008	Temperate
2	Tharandt	50.96	13.57	380	ENF	2003	Temperate
3	Bondville	40.01	−88.29	219	CRO	2000	Temperate
4	Fort Peck	48.31	−105.10	634	GRA	2004	Temperate
5	Morgan Monroe	39.32	−86.41	275	DBF	2007	Temperate
6	Tumbarumba	−35.66	148.15	1200	EBF	2008	Temperate
7	Kaamanen	69.14	27.29	155	TUN	2002	Boreal
8	Hyytiala	61.85	24.29	181	ENF	2003	Boreal
9	Santarem KM67	−2.86	−54.96	130	EBF	2003	Tropical
10	Santarem KM83	−3.02	−54.98	130	EBF	2001	Tropical
11	El Saler	39.35	−0.32	10	ENF	2003	Mediterranean
12	Vaira Ranch	38.41	−120.95	129	GRA	2005	Mediterranean

Table 3. Types of model simulations performed in this study.

	Model simulations ^a	Parameter sets	Meteorological forcing	LAI ^b	Phenology ^c
	local-F	local	FLUXNET	local	Default
local	local-WEIG	local	WFDEI-GPCC	local	Default
vs. global	global-WEIG	global	WFDEI-GPCC	global	Default
data	global-WEIC	global	WFDEI-CRU	global	Default
	global-P	global	PRINCETON	global	Default
Satellite phenology	local-FNM	local	FLUXNET	Site max. MODIS LAI	Default
	local-FM	local	FLUXNET	Site max. MODIS LAI	Daily forcing

^a For model simulation names, local and global refer to the parameter set and F, WEIG, WEIC and P refer to the meteorological forcing data set used.

^b For LAI, local refers to the observed annual maximum LAI at each site and global refers to that obtained from the look-up tables used by the global operational version of the model.

^c Default refers to the default phenology model used by JULES and daily forcing means that the default phenology has been switched off and the model forced with daily MODIS LAI.

2.2 Experimental design

Offline single point simulations of GPP were performed at each of the 12 flux tower sites using various global and local data sets (Table 3). Correct simulation of GPP is important since errors in its calculation can propagate through the model and affect biomass and other flux calculations, such as net ecosystem exchange (NEE) (Schaefer et al., 2012). In JULES, NEE is not a model output and is calculated as total ecosystem respiration minus GPP. The correct representation of leaf level stomatal conductance has an influence on GPP and transpiration. Errors in GPP can also introduce errors into simulated latent and sensible heat fluxes. These study sites (Blyth et al., 2011; Abramowitz et al., 2008, Table 5) were chosen to validate model performance in carbon flux simulation since gap-filled meteorological data, local observations of vegetation and soil characteristics and observed GPP fluxes were available.

One year model simulations were performed and span a range of years due to limited availability of local gap-filled meteorological data, observations of GPP fluxes and vegetation characteristics (Table 2). Prior to performing the model simulations, the soil carbon pools at each site were brought to equilibrium using a 10 year spin-up by cycling 5 year averaged meteorological data (in equilibrium mode), followed by a 1000 year spin-up by cycling observed meteorological data (in dynamical mode). At Tumbarumba, Santarem Km67 and Santarem Km83, 3 year averaged meteorological data was used in the first part of the spin-up process due to limited data availability. More information on model spin-up can be found in Clark et al. (2011).

2.3 Data

JULES requires meteorological data at 6-hourly intervals or less in order to drive the model offline. In this study, half-hourly/hourly meteorological data was used for model

runs using local data and 3-hourly data for simulations using global data. For offline simulations, the model requires downward short-wave and long-wave radiation (W m^{-2}), rainfall and snowfall rate ($\text{kg m}^{-2} \text{s}^{-1}$), air temperature (K), wind speed (m s^{-1}), surface pressure (Pa) and specific humidity (kg kg^{-1}) (Table 1). Gap-filled meteorological forcing data at the local scale was obtained from the FLUXNET network and data at the global scale was obtained from two gridded data sets: WFDEI (Weedon et al., 2014, 2011) and that developed by Sheffield et al. (2006) (referred to as PRINCETON).

Vegetation and soil parameters (Table 1) were adjusted to local or global values depending on the model simulations (Table 3) performed at the 12 flux tower sites. Local vegetation (Tables 5, 6) and soil parameters (not shown) were obtained from the research literature, communications with site Primary Investigator and the Ameriflux data archive. Global vegetation (Tables 5, 6) and soil parameters (not shown) were taken from data sets used in the global operational version of JULES as used in the Hadley Centre Global Environmental Model (HadGEM) climate model. These data sets include the Global Land Cover Characterization (version 2) database (<http://edc2.usgs.gov/glcc/glcc.php>) (PFT fractions), and the Harmonized World Soil Database (version 1.2) (Nachtergaele et al., 2012) (soil texture fractions).

There are several global LAI data sets available, such as ECOCLIMAP (1992) (Masson et al., 2003), CYCLOPES (1997–2007) (Baret et al., 2007), GLOBCARBON (1998–2003) (Deng et al., 2006), MOD15 (2000–present) (Yang et al., 2006) and MISR LAI (2000–present) (Diner et al., 2008; Hu et al., 2007). For the majority of sites used in this study, gap-filled meteorological data and GPP flux observations are only available for the 2000s and therefore, a global data set of satellite LAI was required that covered this period. We used the MODIS LAI product because it is a high spatial and temporal resolution data set with global coverage.

2.3.1 Forcing data

FLUXNET, a “network of regional networks”, is a global network of micrometeorological tower sites that measure the exchange of carbon dioxide, water vapour and energy between the biosphere and atmosphere across a range of biomes and timescales (Baldocchi et al., 2001). Data and site information are available online at: <http://www.fluxnet.ornl.gov/>. Over 500 tower sites are located worldwide on five continents and are used to study a range of vegetation types, such as temperate conifer and broadleaved (deciduous and evergreen) forests, tropical and boreal forests, crops, grasslands, wetlands, and tundra (Baldocchi et al., 2001).

The WATCH Forcing Data (WFD) (1901–2001) was created in the framework of the Water and Global Change (WATCH) project (<http://www.eu-watch.org/>), which sought to assess the terrestrial water cycle using land surface models and general hydrological models. WFD was derived using

the 40 years ECMWF Re-Analysis (ERA-40) for 1958–2001 and data for 1901–1957 was obtained using random years extracted from the ERA-40 data (Weedon et al., 2011). WFD was extended by applying the WFD methodology to the ERA-Interim data for the 1979–2009 period (WFDEI) (Weedon et al., 2014). Within WFD and WFDEI, there are two precipitation products: the first corrected using the Climate Research Unit at the University of East Anglia (CRU) observations; and the second using Global Precipitation Climatology Centre (GPCC) observations. The WFDEI data sets incorporating the GPCC- and CRU-corrected precipitation products are referred to as WFDEI-GPCC and WFDEI-CRU, respectively. WFDEI is only available for land points including Antarctica, and consists of 3-hourly, regularly (latitude–longitude) gridded data at half-degree ($0.5^\circ \times 0.5^\circ$) resolution. This resolution produces a global grid of 360×720 grid cells and is equivalent to a surface resolution of about $56 \text{ km} \times 56 \text{ km}$ at the Equator and $56 \text{ km} \times 32 \text{ km}$ at 55° N (temperate regions).

The Sheffield et al. (2006) data set (PRINCETON) is a global 60-year meteorological data set for driving land surface models developed by the Land Surface Hydrology Research Group at Princeton University. PRINCETON is only available for land points (excluding Antarctica), and consists of 3-hourly, 1° resolution, meteorological data for the 1948–2008 period. This data set has a resolution half that of WFDEI with a global grid of 180×360 grid cells and is equivalent to a surface resolution of about $111 \text{ km} \times 111 \text{ km}$ at the Equator and $111 \text{ km} \times 64 \text{ km}$ at 55° N . The resolution (both spatial and temporal) of the meteorological data can affect the output of land surface and atmospheric chemistry models (Pugh et al., 2013; Ashworth et al., 2010; Ito et al., 2009; Guenther et al., 2006) and may introduce a systematic bias.

2.3.2 Observational data

Local observations of GPP were obtained from the FLUXNET network. Flux tower sites use the eddy covariance method to measure net ecosystem exchange (NEE), which is defined as the net flux of CO_2 , and is separated into GPP and ecosystem respiration with a “flux-partitioning algorithm” (Reichstein et al., 2005). There are a number of approaches used to separate NEE into its two component fluxes, which include extrapolating night-time respiration measurements to the daytime and fitting light-response curves to daytime NEE measurements (Lasslop et al., 2010). In addition to flux-partitioning, the data must also be gap-filled due to unfavourable meteorological conditions and instrument failure (Reichstein et al., 2005). These processes carry with them some uncertainty, which must be quantified. Hagen et al. (2006) found that the uncertainty at the half-hourly timescale was of the order of the observations themselves (i.e. $\sim 100\%$), but only $\sim 10\%$ at the annual timescales for a temperate deciduous forest.

Table 4. Definition of qualitative terms used to describe JULES' ability to simulate GPP when compared to observed FLUXNET GPP. Both RMSE and Bias have units of $\text{g C m}^{-2} \text{ day}^{-1}$. Starting at "Very well", the term associated with the first condition satisfied is used to describe model performance.

Qualitative term	RMSE (= x)	Bias (= y)
<i>Very well</i>	$0 < x < +2$	$ y \leq +1$
<i>Good</i>	$0 < x < +3$	$ y \leq +2$
<i>Poorly</i>	$0 < x < +5$	$ y \geq +2$

2.3.3 Ecological and soil data

The Global Land Cover Characterization (version 2) database, generated by the US Geological Survey, the University of Nebraska-Lincoln, and the European Commission's Joint Research Centre, is a 1 km resolution global land cover data set for use in environmental and modelling research (Loveland et al., 2000). Land cover is classified into 17 categories using the International Geosphere–Biosphere Programme (IGBP) scheme. The land cover category for each of the flux tower sites was extracted from the GLCC database (IGBP code in Table 5). These IGBP codes are then used to derive the annual maximum LAI and canopy height for each PFT from the look-up tables used in the global operational version of JULES. Further information on how these variables are derived can be found in Appendix A.

Global soil texture fractions (% of sand, silt and clay) for each of the 12 FLUXNET sites (not shown here) were extracted from the Harmonized World Soil Database (version 1.2) (HWSD) (Nachtergaele et al., 2012). The equations used to compute soil hydraulic and thermal characteristics were taken from the Unified Model Documentation Paper No. 70 (Jones, 2007). Note that the equations in Jones (2007) apply only to mineral soils, as organic soils behave differently (Gornall et al., 2007). In this study, the soils are classified as mineral at all 12 sites. Since the HWSD contains soil textures for two soil depths (0–30 and 30–100 cm) and JULES contains four soil layers (thicknesses of 0.1, 0.25, 0.65 and 2.0), the 0–30 cm soil textures were assigned to the top two model soil layers (thicknesses 0.1 and 0.25 m, respectively), and the 30–100 cm textures were assigned to the bottom two layers (thicknesses 0.65 and 2.0 m, respectively). The local soil textures are provided as site averages and therefore, each model soil layer (four in total) is assigned the same set of soil textures.

2.3.4 MODIS LAI products

The MODIS LAI product, computed from MODIS spectral reflectances, provides continuous and consistent LAI coverage for the entire global land surface at 1 km resolution (Yang et al., 2006). Some gaps and noise in the data are

possible due to the presence of cloudiness, seasonal snow cover and instrument problems, and this can limit the usefulness of the product (Gao et al., 2008; Lawrence and Chase, 2007). In this study, we use the MODIS Land Product Subsets, created by the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC), which provide summaries of selected MODIS Land Products for use in model validation and field site characterisation and include data for more than 1000 field sites and flux towers (<http://daac.ornl.gov/MODIS/>).

The MODIS Land Product Subsets (ASCII format) contain LAI data for a $7 \text{ km} \times 7 \text{ km}$ grid of 49 pixels, with each pixel representing the $1 \text{ km} \times 1 \text{ km}$ scale, at 8-day composite intervals. The average of the 3×3 pixel grid box centred on the flux tower is taken to be that day's LAI value. Only pixel values with an even quality control (QC) flag was used for the averaging and this produced a time-series of 8-day observations at each of the sites. Missing data were dealt with by using the previous good value in the time-series. The exception to this was Bondville, where missing data occurred in January 2000, since MODIS only started recording data in February 2000 (this year was used due to limited data availability at the site). To gap-fill the missing data, an 11-year average was computed and the missing data replaced with the average for January 2000. Finally, each time-series of 8-day composite values was linearly interpolated to obtain a daily LAI time-series.

2.4 Outline of experiments

This section describes the model simulations performed in the study. In the model simulation names, local and global refer to the parameter set and F, WEIG, WEIC and P refer to the meteorological forcing data set used (Table 3). Vegetation competition has been switched off for all model simulations.

2.4.1 Effect of local data on simulated GPP

Using JULES3.0, we compared model simulations using local parameter and meteorological data sets (local-F; Table 3) to observations of GPP from the FLUXNET network. For this set of model simulations, the default phenology model (used to update LAI) and TRIFFID were used.

The ability of the model to simulate interannual variability was also examined. Multi-year model simulations were performed for six of the sites using local data; one from each of the various climate zones (Harvard Forest, Vaira Ranch, Hyytiälä, Santarem Km67), the Southern Hemisphere site (Tumbarumba) and the temperate site, Morgan Monroe. Since meteorological data was available for multiple years at these sites, but not model parameter data, the same parameter data sets used for the single-year runs (Table 2) would be used for the multi-year runs at specific sites.

Table 5. Vegetation (PFT) and non-vegetation land cover type (BL: broadleaf tree, NL: needleleaf tree, C3g: C3 grass, C4g: C4 grass, sh: shrubs, bs: bare soil) fractions at the 12 FLUXNET sites. For each site, the first row refers to global data and the second refers to local.

Site	IGBP code	IGBP class	Plant functional types					References	
			BL	NL	C3g	C4g	sh		bs
Harvard Forest	4	DB forest	0.60		0.05	0.10	0.05	0.20	Urbanski et al. (2007)
		DB forest	0.95					0.05	
Vaira Ranch	8	Woody savannah	0.50		0.15		0.25	0.10	Ryu et al. (2008)
		Grassland			0.95			0.05	
Morgan Monroe	4	DB forest	0.60		0.05	0.10	0.05	0.20	Schmid et al. (2000)
		DB forest	0.90					0.10	
Hyytiala	1	EN forest		0.70	0.20			0.10	Suni et al. (2003)
		EN forest		0.95				0.05	
Tharandt	5	Mixed forest	0.35	0.35	0.20			0.10	Grünwald and Bernhofer (2007)
		EN forest		0.95				0.05	
Tumbarumba	2	EB forest	0.85			0.10		0.05	Leuning et al. (2005)
		EN forest		0.90				0.10	
El Saler	7	Open shrub			0.05	0.10	0.35	0.50	Stöckli et al. (2008)
		EN forest		0.90				0.10	
Fort Peck	10	Grassland			0.70	0.15	0.05	0.10	Gilmanov et al. (2005)
		Grassland			0.90			0.10	
Kaamanen	1	EN forest		0.70	0.20			0.10	Laurila et al. (2001)
		Grassland			0.90			0.10	
Santarem KM67	2	EB forest	0.85			0.10		0.05	Hutyra et al. (2007)
		EB forest	0.98					0.02	
Santarem KM83	2	EB forest	0.85			0.10		0.05	Goulden et al. (2004)
		EB forest	0.98					0.02	
Bondville	12	Cropland			0.75	0.05		0.20	Meyers and Hollinger (2004)
		Grassland			0.90			0.10	

2.4.2 Effect of global data on simulated GPP

Using JULES3.0, we compared model simulations using parameter sets from the HadGEM model and global meteorological data (global-WEIG, global-WEIC and global-P; Table 3) to observations of GPP from the FLUXNET network. In addition to this, we quantified how much error is introduced into model simulations (using local model parameters) when using global (WFDEI-GPCC) instead of local meteorological data (local-WEIG and local-F in Table 3). In these model simulations, the default phenology model and TRIFID were used.

2.4.3 Comparison of global to local meteorological data

The WFDEI-GPCC, WFDEI-CRU and PRINCETON data sets were compared to FLUXNET to find out which one more closely captures the local meteorological conditions.

2.4.4 Daily satellite phenology

Using JULES3.0 and JULESmod, we tested the ability of the JULES phenology model to simulate the seasonal cycle of GPP by comparing model simulations, where JULES uses MODIS LAI data (local-FM and local-FNM; Table 3) to those using the default phenology model (local-F; Table 3). When using the default phenology module, LAI was computed internally by scaling the annual maximum LAI, which was then used to calculate GPP. When forcing JULES with daily MODIS data (local-FM), the phenology module was switched off and the MODIS LAI then used to compute GPP. For model simulations using MODIS data and the default phenology module (local-FNM), the annual maximum MODIS LAI is set to be the annual maximum LAI. Vegetation competition has been switched off and local parameters used for both sets of model simulations (local-FM, local-FNM).

Table 6. Local and global biophysical parameters (site annual maximum LAI, canopy height and V_{cmax}) at the 12 FLUXNET sites. For each site, the first row refers to global data, the second refers to local and the third refers to satellite. Online data was accessed in April 2013.

Site	LAI ($\text{m}^2 \text{m}^{-2}$)	Canopy height (m)	V_{cmax} ($\mu\text{mol CO}_2 \text{m}^{-2} \text{s}^{-1}$)	References
Harvard Forest	5.00	19.01	32.00	^a Harvard Forest data archive/exchange
	5.00 ^a	24.00 ^a	35.20 ^a	
	6.03	–	–	
Vaira Ranch	4.00	1.26	48.00	^b Ameriflux biological data ^c Beerling and Quick (1995)
	2.74 ^b	0.67 ^b	42.25 ^c	
	3.46	–	–	
Morgan Monroe	5.00	19.01	32.00	
	5.23 ^b	27.00 ^b	34.80 ^c	
	6.81	–	–	
Hyttiala	6.00	21.46	24.00	^d P. Kolari (personal communication, 2013) ^e Suni et al. (2003)
	3.00 ^d	14.00 ^e	60.00 ^d	
	4.56	–	–	
Tharandt	6.00	21.46	24.00	^f T. Grünwald (personal communication, 2013) ^g Grünwald and Bernhofer (2007) ^h Kattge et al. (2009)
	7.10 ^f	26.50 ^g	62.50 ^h	
	3.82	–	–	
Tumbarumba	4.00	16.38	24.00	ⁱ E. van Gorsel (personal communication, 2013) ^j Cleugh et al. (2007) ^k Haverd et al. (2009)
	2.50 ⁱ	40.00 ^j	74.33 ^k	
	6.08	–	–	
El Saler	4.00	16.38	24.00	^l Blyth et al. (2010) ^m Obtained from http://www.bgc-jena.mpg.de
	4.00 ^l	12.00 ^m	62.5 ^h	
	1.04	–	–	
Fort Peck	3.00	1.04	48.00	ⁿ Obtained from: http://ameriflux.ornl.gov/
	2.00 ^l	0.40 ⁿ	42.25 ^s	
	1.41	–	–	
Kaamanen	2.00	0.79	48.00	^o Laurila et al. (2001) ^p Aurela et al. (1998)
	0.70 ^o	1.00 ^p	42.25 ^c	
	1.33	–	–	
Santarem Km67	9.00	28.12	32.00	^q Oak Ridge National Laboratory DAAC ^r Hutyra et al. (2007) ^s Domingues et al. (2007)
	5.25 ^q	45.00 ^r	81.00 ^s	
	6.73	–	–	
Santarem Km83	9.00	28.12	32.00	^t Doughty and Goulden (2008) ^u Bruno et al. (2006) ^v Domingues et al. (2007)
	6.00 ^t	40.00 ^u	81.00 ^v	
	6.63	–	–	
Bondville	5.00	1.46	48.00	^w Meyers and Hollinger (2004)
	6.74 ^b	0.90 ^w	117.35 ^c	
	3.37	–	–	

2.5 Model analyses

To quantify differences between output from the various model simulations and observations, we used root mean squared error (RMSE) (Eq. 1), which is a measure of the average error of the simulations, and bias (Eq. 2), which is the average difference between model and observations (a measure of under- or overprediction) with the absolute (Eq. 3)

and percentage differences (Eq. 4):

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{t=n} (x_t - x_{o,t})^2}{n}}, \quad (1)$$

$$\text{Bias} = \frac{\sum_{t=1}^{t=n} x_t - \sum_{t=1}^{t=n} x_{o,t}}{n}. \quad (2)$$

x_t and $x_{o,t}$ are model and observed daily GPP fluxes, respectively, which have been smoothed using a 7-day moving aver-

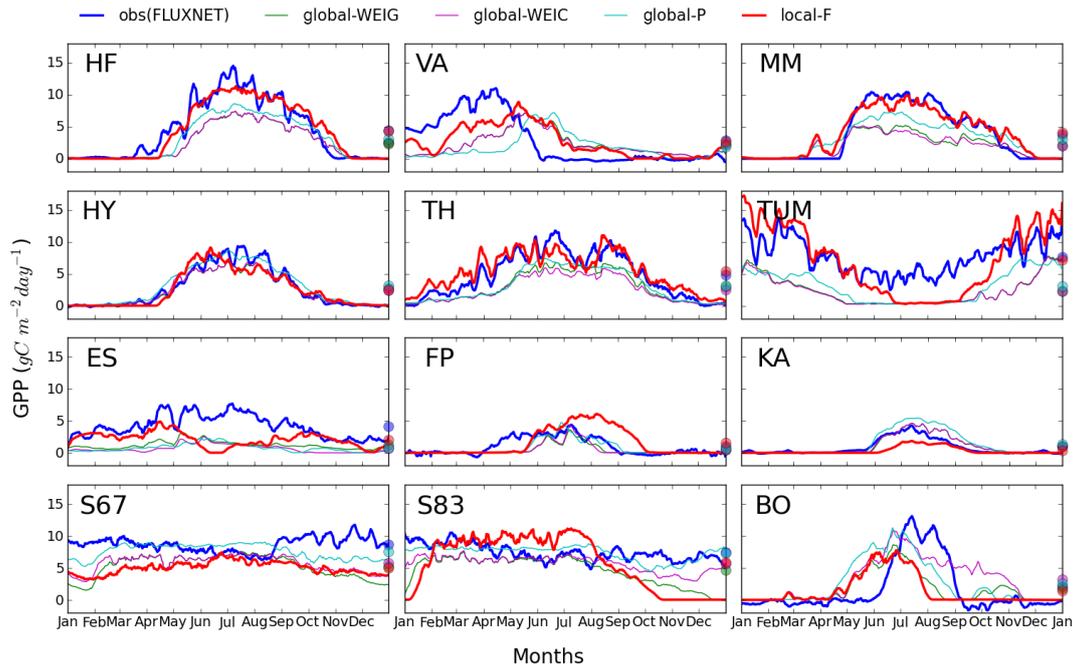


Figure 1. Seasonal cycle of model-predicted (local-F, global-WEIG, global-WEIC and global-P in Table 3) and observed GPP fluxes, smoothed with a 7-day moving average window, at the 12 FLUXNET sites (HF: Harvard Forest; VA: Vaira Ranch; MM: Morgan Monroe; HY: Hyytiala; TH: Tharandt; TUM: Tumberumba; ES: El Saler; FP: Fort Peck; KA: Kaamanen; S67: Santarem Km67; S83: Santarem Km83; BO: Bondville). Model simulation years are given in Table 2. The thick lines refer to FLUXNET observations (blue) and simulated GPP from local-F model simulations (red). Annual averages for model simulations and observations are plotted as thick dots on the right of each plot in the same colours.

age since we are interested in the long-term average and not daily variability. n is the number of paired values (number of days in year). The absolute difference (ΔGPP) between the model and observations is the absolute value of the difference in total annual GPP for each and the percentage difference ($\Delta\%$) is the absolute difference divided by the observed total annual GPP:

$$\Delta\text{GPP} = \left| \sum \text{GPP}_{\text{obs}} - \sum \text{GPP}_{\text{model}} \right|, \quad (3)$$

$$\Delta\% = \frac{\Delta\text{GPP}}{\sum \text{GPP}_{\text{obs}}} \times 100. \quad (4)$$

In order to describe JULES' ability to reproduce simulated GPP, a simple, but subjective, ranking system using qualitative terms (*Very well*, *Good* and *Poorly*) was devised based on RMSE and bias (Table 4, Fig. 2a). These ranges were used as interannual variability was about $\pm 1 \text{ gC m}^{-2} \text{ day}^{-1}$ in both RMSE and bias (Sect. 3.1).

3 Results

3.1 Effect of local data on simulated GPP

When driven with local meteorological and parameter data sets (local-F; Fig. 1), JULES has a negative bias with total annual GPP underestimated by 16% ($3049 \text{ gC m}^{-2} \text{ year}^{-1}$;

Table 7) across all sites compared to observations. By using local data, JULES performs *very well* (see Fig. 2a and Table 4 for definition of qualitative terms used to describe model performance) at the temperate forest sites, Harvard Forest, Morgan Monroe, Hyytiala and Tharandt, where RMSEs range from $1.1\text{--}1.4 \text{ gC m}^{-2} \text{ day}^{-1}$, biases from -0.2 to $+0.3 \text{ gC m}^{-2} \text{ day}^{-1}$ (Fig. 2a) and absolute differences from $40\text{--}211 \text{ gC m}^{-2} \text{ year}^{-1}$ (Table 7) and *good* at Vaira Ranch with an RMSE of $2.78 \text{ gC m}^{-2} \text{ day}^{-1}$, bias of $-0.19 \text{ gC m}^{-2} \text{ day}^{-1}$ and absolute difference of $71 \text{ gC m}^{-2} \text{ year}^{-1}$. The model performs *poorly* at Tumberumba, El Saler, Bondville and the tropical sites, Santarem Km67 and Santarem Km83, with RMSEs ranging from $1.8\text{--}4.1 \text{ gC m}^{-2} \text{ day}^{-1}$, biases from -3.7 to $-0.2 \text{ gC m}^{-2} \text{ day}^{-1}$ and absolute differences from $71\text{--}1340 \text{ gC m}^{-2} \text{ year}^{-1}$.

At the temperate forest sites, JULES simulates the summer carbon uptake and leaf onset and senescence *very well*. For example, at the needleleaf forests, Hyytiala and Tharandt, the model correctly captures the timing and magnitude of the seasonal cycle of GPP (Fig. 1). JULES is able to capture the beginning and ending of the growing season, but underestimates the summer carbon uptake at Tumberumba, a temperate sclerophyll forest (forests dominated by plants that have hard leaves and are adapted to drought) (Fig. 1). At the tropical sites, Santarem Km67 and Santarem Km83, the seasonal cycle has been modelled *poorly*, with the total annual

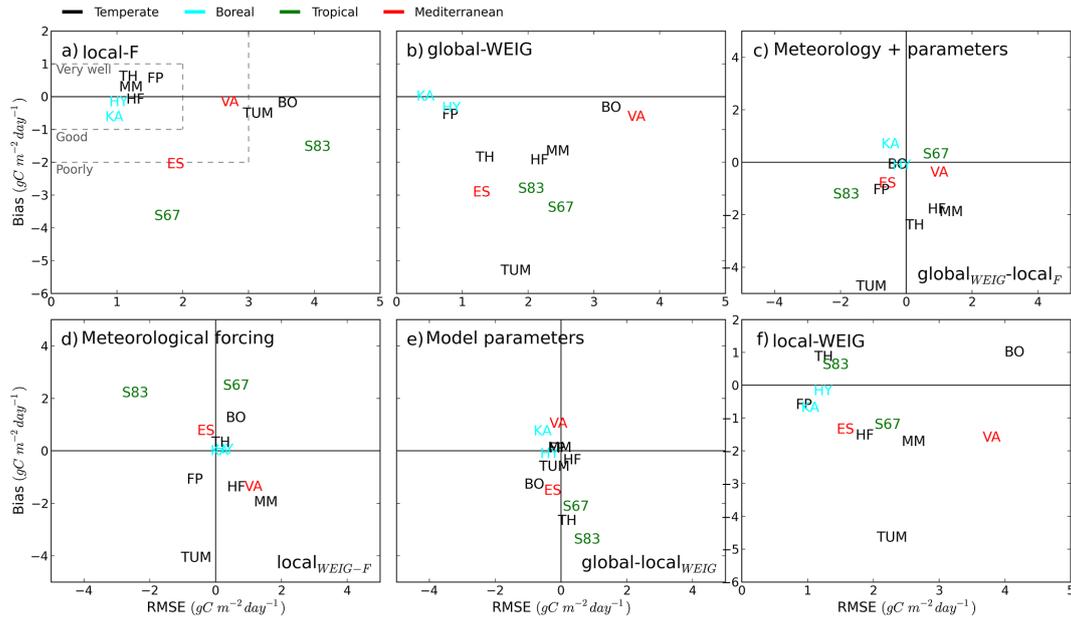


Figure 2. Comparison of modelled and observed GPP using bias and RMSE at the 12 FLUXNET sites (HF: Harvard Forest; VA: Vaira Ranch; MM: Morgan Monroe; HY: Hyytiala; TH: Tharandt; TUM: Tumberumba; ES: El Saler; FP: Fort Peck; KA: Kaamanen; S67: Santarem Km67; S83: Santarem Km83; BO: Bondville) for three sets of model simulations: **(a)** local-F; **(b)** global-WEIG; and **(f)** local-WEIG (Table 3). **(c)** displays the differences between bias and RMSE for global-WEIG and local-F model simulations; **(d)** differences between local-WEIG and local-F model simulations; and **(e)** differences between global-WEIG and local-WEIG model simulations. Marked on **(c)**, **(d)** and **(e)** next to the figure letter are how the sets of model simulations differ. The site labels are coloured according to their climate zone (Table 2). The dashed lines on **(a)** show the regions defined by the qualitative terms (Table 4) used to describe model performance.

Table 7. Absolute and percentage differences between model simulated and observed (FLUXNET) total annual GPP ($\text{gC m}^{-2} \text{ year}^{-1}$) at the 12 flux tower sites. $\sum \text{GPP}_{\text{obs}}$ is the observed total annual GPP, ΔGPP is the absolute difference (Eq. 3) between the model and observed total annual GPP, and $\Delta \%$ is the percentage difference (Eq. 4) between the model and observed total annual GPP. Values highlighted in bold mean that the difference is negative (i.e. $\sum \text{GPP}_{\text{obs}} < \sum \text{GPP}_{\text{model}}$). The total value for each of the model simulations was computed using the differences and not the absolute differences.

Site	FLUXNET	local-F		local-WEIG		global-WEIG		global-WEIG		global-P	
	$\sum \text{GPP}_{\text{obs}}$	ΔGPP	$\Delta \%$								
Harvard Forest	1621	40	2	567	35	716	44	711	44	486	30
Vaira Ranch	1047	71	7	592	57	235	22	259	25	369	35
Morgan Monroe	1385	94	7	639	46	616	44	661	48	256	18
Hyytiala	997	68	7	73	7	135	14	120	12	144	14
Tharandt	1754	211	12	306	17	687	39	819	47	590	34
Tumberumba	2806	197	7	1710	61	1951	70	1984	71	1690	60
El Saler	1512	760	50	499	33	1073	71	1276	84	1234	82
Fort Peck	367	194	53	229	62	213	58	200	54	105	29
Kaamanen	368	249	68	273	74	8	2	5	1	124	34
Santarem Km67	3171	1340	42	451	14	1245	39	1075	34	392	12
Santarem Km83	2724	583	21	202	7	1033	38	644	24	40	1
Bondville	766	240	31	200	26	131	17	406	53	177	23
Total	18 518	3049		4325		8043		7348		4717	

GPP being underestimated by 42 % ($1340 \text{ gC m}^{-2} \text{ year}^{-1}$) and 21 % ($583 \text{ gC m}^{-2} \text{ year}^{-1}$), respectively (Table 7).

JULES can simulate interannual variability when using local data with average RMSEs across all six sites for all years

being within $0.7 \text{ gC m}^{-2} \text{ day}^{-1}$ and average biases within $1.2 \text{ gC m}^{-2} \text{ day}^{-1}$ of model results from the corresponding single-site runs (Fig. 3). Interannual variability is captured *very well* at the temperate sites (Harvard Forest, Hyytiala

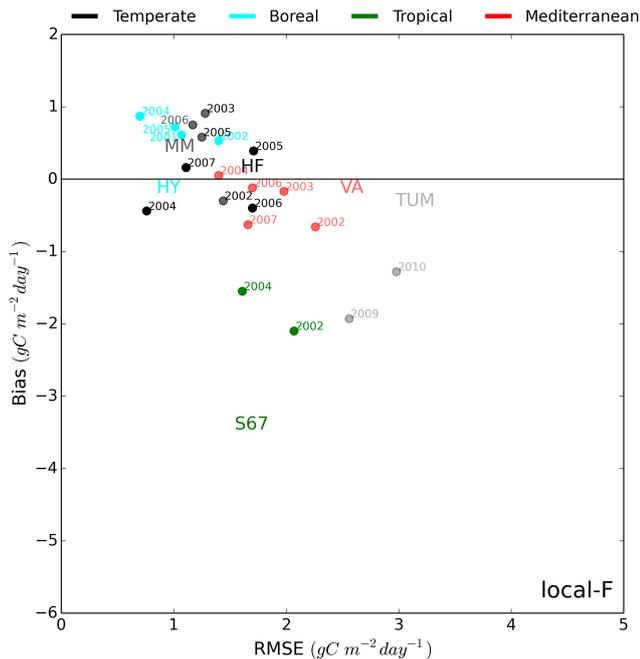


Figure 3. Multi-year comparison of modelled and observed GPP using bias and RMSE at six FLUXNET sites (HF: Harvard Forest; VA: Vaira Ranch; MM: Morgan Monroe; HY: Hyytiala; TUM: Tumbarumba; S67: Santarem Km67) for model simulations using local parameter and meteorological data (local-F). The site labels are coloured according to their climate zone (Table 2) and represent data from model simulations performed for the year specified in Table 2, with results from other years plotted using the model simulation year and labels coloured the same as the original site label.

and Morgan Monroe) and Vaira Ranch with RMSEs ranging from +1 to +3 $\text{gC m}^{-2} \text{day}^{-1}$ and biases from +1 to $-1 \text{ gC m}^{-2} \text{day}^{-1}$. As observed with the single-site model simulations, the model fails to capture interannual variability at Santarem Km67 and Tumbarumba (Fig. 3).

Overall, JULES performs *very well* with the use of local data (meteorological and parameter data sets) with negative biases observed at the tropical sites and the Southern Hemisphere site, Tumbarumba, with the same trend also observed when the model simulates interannual variability.

3.2 Effect of global data on simulated GPP

By replacing the local data with global parameter and meteorological data, JULES had a much greater negative bias with total annual GPP underestimated by 30% ($6703 \text{ gC m}^{-2} \text{year}^{-1}$; Table 7) on average across all sites compared to observations (global-WEIG, global-WEIC and global-P; Fig. 1). This is also shown in the annual average GPP, which has been plotted for each of the model simulations and observations at the 12 sites (Fig. 1) and the percentage differences (Table 7), which are, in general, larger for simulations using global data than for those using local. This

trend occurs at all sites, with the exception of the wetland site, Kaamanen, and Santarem Km83, where modelled total annual GPP ($2684 \text{ gC m}^{-2} \text{year}^{-1}$ and $492 \text{ gC m}^{-2} \text{year}^{-1}$, respectively) is overestimated (global-P; Table 7) compared to model runs using only local data ($2141 \text{ gC m}^{-2} \text{year}^{-1}$ and $119 \text{ gC m}^{-2} \text{year}^{-1}$, respectively; Table 7).

As well as quantifying differences in model simulations using either local or global data, it is useful to know how global meteorological data affects local model runs. Global meteorological data can be used in place of FLUXNET data in order to drive JULES (local-WEIG; Table 3). This is important for ecological research sites where there is limited or no local meteorological data available. Using the WFDEI-GPCC meteorological data set (local-WEIG; Table 3) to force the model increases the negative bias of model simulations using only local data (Fig. 2f), with a 7% reduction in simulated total annual GPP ($15469 \text{ gC m}^{-2} \text{year}^{-1}$ for local-F reduced to $14193 \text{ gC m}^{-2} \text{year}^{-1}$ for local-WEIG; Table 7).

Forcing the model with WFDEI-GPCC (local-WEIG) results in decreases in model performance (increases in bias and RMSE) at the majority of sites. The tropical sites, Santarem Km67 and Santarem Km83, are two exceptions and show a noticeable improvement in modelled yearly GPP (66 and 61% reduction of bias, respectively) and changes to modelled seasonal cycle (25% increase and 65% reduction of RMSE, respectively). However, at some sites, such as Tharandt, Kaamanen and Hyytiala, forcing JULES with global meteorological data has not introduced large negative biases into GPP predictions (Table 7), with RMSEs ranging from $1.1\text{--}1.3 \text{ gC m}^{-2} \text{year}^{-1}$ (Fig. 2f).

In general, we found the meteorological data had a greater impact on modelled GPP fluxes than model parameters. Larger differences exist between local-WEIG and local-F (local_{WEIG-F}; Fig. 2d), which differ only in the atmospheric forcings data set used, compared to between global-WEIG and local-WEIG (global – local_{WEIG}; Fig. 2e), which differ only in the model parameter sets used.

The ability of JULES to capture yearly GPP (bias) and the seasonal cycle (RMSE) is affected at the majority of sites when using global meteorological data (Fig. 2d), with improvements observed at Santarem Km67 and Santarem Km83. However, model parameters were found to affect bias at all 12 sites (Fig. 2e) with the tropical sites being the most influenced. With the exception of Tumbarumba, biases associated with meteorological data compensate for those associated with model parameters at the tropical sites (global_{WEIG} – local_F; Fig. 2c).

Overall, we found that with the use of global data (model parameter and meteorological data), model performance decreased from *very well* to *good* and *poorly* at most sites, with the exception of the tropical sites. Driving JULES with global meteorological data introduces biases into single site simulations. At the majority of sites, these biases are negative, but at tropical sites, the global meteorological data

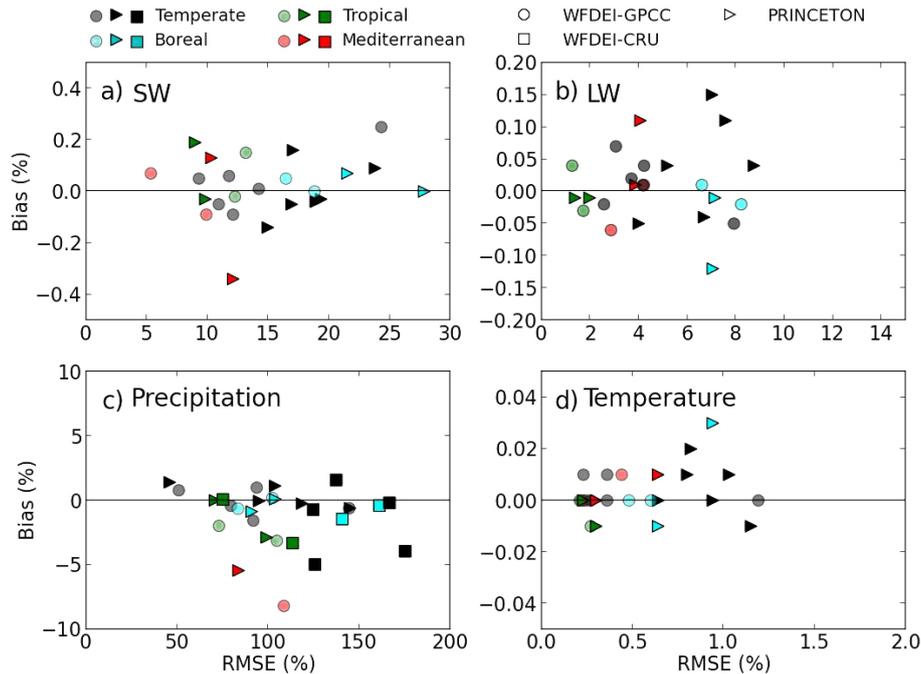


Figure 4. Bias and RMSE, expressed as percentages of daily average, when comparing global (WFDEI-GPCC (circles), WFDEI-CRU (squares) and PRINCETON (triangles)) to local meteorological data for four meteorological variables: (a) downward short-wave radiation (SW); (b) downward long-wave radiation (LW); (c) precipitation; and (d) surface air temperature, at the 12 FLUXNET sites (HF: Harvard Forest; VA: Vaira Ranch; MM: Morgan Monroe; HY: Hyytiala; TH: Tharandt; TUM: Tumbarumba; ES: El Saler; FP: Fort Peck; KA: Kaamanen; S67: Santarem Km67; S83: Santarem Km83; BO: Bondville). The site labels are coloured according to their climate zone (Table 2). Note that before computing bias and RMSE, the meteorological data was normalised against the annual mean for each site.

improves model performance. We found the meteorological data to have a greater impact on GPP fluxes than model parameters.

3.3 Global vs. local meteorological data

As well as quantifying the error introduced into model simulations by using global meteorological data instead of local, we also compared the global meteorological data to local data. Only the downward short-wave and long-wave radiation fluxes, precipitation and surface air temperature variables have been compared to FLUXNET values, since these variables play the most influential role of the meteorological forcings in canopy photosynthesis and light propagation in JULES (Alton et al., 2007). In order to compare the meteorological data sets, the data was normalised against the annual mean for each site before computing the RMSE and bias.

Of the two global meteorological data sets used in this study, the WFDEI data set compares best to FLUXNET (lower RMSEs and biases than PRINCETON) at the majority of sites (Fig. 4). Surface air temperatures compare best to local meteorological measurements with average RMSEs of 0.4 and 0.7% (7 day filtered RMSE expressed as percentages of the annual mean value) (1.5 and 2.4 K) across all sites for the WFDEI and PRINCETON data sets, respec-

tively (Fig. 4d), followed by the downward short-wave radiation fluxes with average RMSEs of 13 and 17% (27.0 and 33.2 W m^{-2}) for WFDEI and PRINCETON, respectively (Fig. 4a), and downward long-wave radiation fluxes with average RMSEs of 4 and 5% (18.9 and 25.0 W m^{-2}) for WFDEI and PRINCETON, respectively (Fig. 4b). Precipitation data from global data sets differ most from local values with RMSEs of 112–178% (2.7–4.4 mm day^{-1}) for WFDEI-GPCC, WFDEI-CRU and PRINCETON, respectively, which may be due to how the precipitation products of each global data set is corrected (Weedon et al., 2011; Sheffield et al., 2006).

In addition to comparing the global meteorological variables to their local values, we also examine the two precipitation products, WFDEI-GPCC (GPCC-corrected) and WFDEI-CRU (CRU-corrected), within the WFDEI data set. We found WFDEI-GPCC and WFDEI-CRU compare equally well at the 12 FLUXNET sites (Fig. 4c) with average RMSEs of 2.7 and 2.8 mm day^{-1} , respectively. Differences between GPCC- and CRU-corrected precipitation RMSEs are small (0.0–1.4 $\text{g C m}^{-2} \text{day}^{-1}$) at individual flux tower sites. When forcing JULES with WFDEI, there is little difference when either WFDEI-GPCC or WFDEI-CRU is used as the precipitation product, with average RMSEs of 2.9 and 2.8 $\text{g C m}^{-2} \text{day}^{-1}$, respectively, across all sites, al-

though differences in the data sets may be more important when JULES is run globally.

Even though WFDEI compares better to the local meteorological data than PRINCETON, we found that when JULES is forced with the PRINCETON data set, improvements in GPP predictions were observed at Santarem Km67 and Santarem Km83 (Fig. 1). We observed that at the tropical sites, the meteorological forcings were the primary driver of productivity for model simulations using global data and that biases associated with the global meteorological data compensated for incorrect parameter values.

By swapping local meteorological data with global meteorological data (PRINCETON) for model simulations using local data (local-F), it was found that the positive bias associated with global surface air temperature (PRINCETON) at Santarem Km83 is the primary cause of improved model performance (39 % reduction in RMSE) when using global data and by forcing JULES with the PRINCETON data set and using the lower global V_{cmax} value (Table 6), the model was able to reproduce the seasonal cycle *very well* (RMSE of $1.26 \text{ g C m}^{-2} \text{ day}^{-1}$). At Santarem Km67, we found the downward long-wave radiation to be the main reason for the improved seasonal cycle (35 % reduction in RMSE) and by using the PRINCETON data set and global V_{cmax} value (Table 6), model performance was improved (RMSE of $2.12 \text{ g C m}^{-2} \text{ day}^{-1}$).

Compensation between meteorological data and model parameters also occurs at Hyytiala, where the model performs *very well* with global meteorological and parameter data sets (Fig. 1). The global downward short-wave radiation is larger than its locally measured value and this offsets the low global V_{cmax} value at this site (Table 6, Fig. 6b).

Overall, we found the WFDEI data set compares better than PRINCETON to FLUXNET and of the four meteorological variables examined, the radiation fluxes (downward short-wave and long-wave) and surface air temperatures compare *very well* to local values. Within the WFDEI data set, the two precipitation products (WFDEI-GPCC and WFDEI-CRU) compare equally well to FLUXNET precipitation. Improvements were observed at the tropical sites when JULES is forced with PRINCETON and this is due to biases associated with the meteorological data.

3.4 Forcing JULES with daily satellite phenology

The performance of LSMs depend on how well the seasonal variation of LAI is represented, since GPP is strongly influenced by the timing of budburst and leaf senescence (Liu et al., 2008). In JULES, LAI is essential for the calculation of plant canopy photosynthesis and is updated daily in response to temperature. We test the JULES phenology model by comparing model predictions of GPP when JULES uses its default phenology model with those in which JULES uses local data with the annual maximum LAI set to be the MODIS annual maximum LAI (local-FNM), and with those in which

the model uses local data and is forced with daily MODIS LAI (local-FM).

Forcing JULES with daily satellite LAI (local-FM) results in either small improvements (average reduction in RMSE by $0.2 \text{ g C m}^{-2} \text{ day}^{-1}$) or none at all at the 12 flux tower sites (Fig. 5c). An average RMSE of $2.2 \text{ g C m}^{-2} \text{ day}^{-1}$ across all sites is observed when the model is forced with daily MODIS LAI (local-FM), which is less than that for model simulations using no MODIS information (local-F; average RMSE of $2.4 \text{ g C m}^{-2} \text{ day}^{-1}$) and those which use the annual maximum MODIS LAI as the annual maximum LAI at each site (local-FNM; average RMSE of $2.39 \text{ g C m}^{-2} \text{ day}^{-1}$).

By using MODIS data, there is only a small reduction (8 and 0.04 % for local-FM and local-FNM, respectively) in average RMSE when simulating GPP compared to model runs that do not use it. Of the 12 sites, only seven (Harvard Forest, Vaira Ranch, Hyytiala, Tharandt, Tumberumba, Kaamanen and Santarem Km67) show improved model performance when either being forced with daily MODIS LAI (Fig. 5c) or using the annual maximum MODIS LAI as the model annual maximum LAI (Fig. 5b). At these seven sites, simulated yearly GPP increases in total by 21 %. At the remaining sites, JULES performs better using the default phenology module (Fig. 5a).

Of the seven sites where JULES' performance improved using MODIS data, forcing JULES with daily satellite phenology (local-FM) only resulted in improved model performance at Santarem Km67 (Fig. 5c) and at the remaining six sites, using the default phenology with the annual maximum MODIS LAI set to be the annual maximum LAI (Fig. 5b), JULES' performance improved. Even with the addition of MODIS data, the model still performed *poorly* at Bondville, with only a slight improvement in predicted GPP (1 and 15 % reduction of RMSE for local-FM and local-FNM, respectively) compared to using only local data (RMSE of $3.66 \text{ g C m}^{-2} \text{ day}^{-1}$).

The sites which display the largest improvements in simulated GPP, when forced with MODIS LAI, are those which have low LAI values (54 and 24 % reduction in RMSE at Vaira Ranch and Fort Peck, respectively) (Fig. 5c). Small improvements were also observed at the tropical sites (13 and 14 % reduction in RMSE at Santarem Km67 and Santarem Km83, respectively). At some sites, using MODIS data had no effect on model results (El Saler) and in some cases, the model performed worse (Tumberumba).

The total annual simulated GPP for model runs using MODIS data ($15,334$ and $15,227 \text{ g C m}^{-2} \text{ year}^{-1}$, for local-MF and local-NMF, respectively) is slightly lower than when using only local data ($15,469 \text{ g C m}^{-2} \text{ year}^{-1}$), but better than when using global data (global-WEIG; $14,193 \text{ g C m}^{-2} \text{ year}^{-1}$). This is a result of the annual maximum MODIS LAI being closer to local values than global (Fig. 5a). The increased LAI of the global data does not result in increased GPP predictions since the meteorological

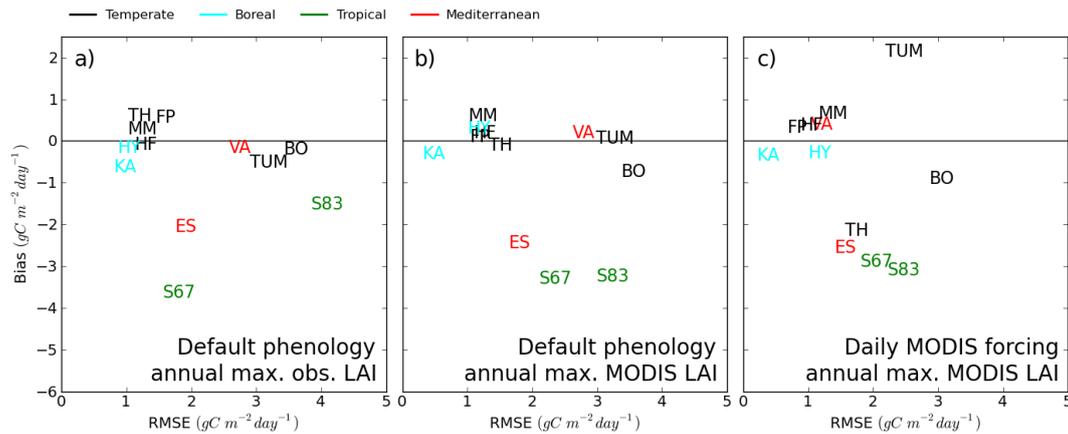


Figure 5. Comparison of modelled and observed GPP using bias and RMSE (computed using anomalies) at the 12 FLUXNET sites (HF: Harvard Forest; VA: Vaira Ranch; MM: Morgan Monroe; HY: Hyytiälä; TH: Tharandt; TUM: Tumberumba; ES: El Saler; FP: Fort Peck; KA: Kaamanen; S67: Santarem Km67; S83: Santarem Km83; BO: Bondville) for three sets of model simulations: (a) default phenology model with locally observed annual maximum LAI (data values used same as in Fig. 2a (local-F)); (b) default phenology model with annual maximum MODIS LAI (model simulations local-FNM); and (c) daily MODIS forced model simulations with annual maximum MODIS LAI (model simulations local-FM). The site labels are coloured according to their climate zone (Table 2).

data and vegetation parameters, such as V_{cmax} , may have a greater impact on predicted GPP than LAI.

Overall, when JULES is forced with daily MODIS LAI small improvements (8 % reduction in average RMSE; local-FM) in predicted GPP are observed at a number of sites, though there exists a negative bias associated with using MODIS data. By setting the annual maximum MODIS LAI to be the annual maximum LAI at each site, the model performs equally well (0.04 % reduction in average RMSE; local-FNM) to local model simulations. We also observed improvements in simulated GPP at sites with low LAI values, such as grasslands, when JULES is forced with daily LAI.

4 Discussion

4.1 How well does JULES perform when using the best available local meteorological and parameter data sets compared to those using global data?

At more than half of the sites, JULES performs *very well* when using local meteorological and parameter data sets with a negative bias observed for the remaining sites (Fig. 2a). At the six sites where multi-year model simulations were performed, interannual variability is captured by the model using local data with the exception of Santarem Km67 and Tumberumba. This trend is also observed with the single-year runs.

The use of global parameter and meteorological data sets introduces a negative bias into GPP simulations at all sites with the exception of the mediterranean site, El Saler, and the tropical sites (Fig. 2b). Using local parameter and global meteorological data to drive JULES (local-WEIG) increases the

negative bias of local model simulations (local-F) (Fig. 2f). We observed decreases in model performance at the majority of sites, with the exceptions being the tropical sites (Santarem Km67/Km83). At some sites, such as Hyytiälä and Kaamanen, using global meteorological data produced similar results (Fig. 2a, f) to using FLUXNET data.

Our results compare well with the evaluation of JULES by Blyth et al. (2011), where parameters were obtained as though the model was embedded in a GCM. Differences between the two studies include different model versions and global meteorological data sets used. Comparing our results with Fig. 3 of Blyth et al. (2011), we also found simulated photosynthesis to be underestimated for the temperate forests (Harvard Forest, Tharandt and Morgan Monroe), grasslands (Fort Peck), mediterranean sites (El Saler) and the tropical forests (Santarem Km67), and overestimated for the wetlands (Kaamanen). We observed that the use of local observations of site characteristics, such as PFT fractions and vegetation properties, lead to improvements in model performance at more than half of the sites (Fig. 2a), though errors still exist with percentage differences ranging from 2–12 %.

Differences between global and local data include PFT fractions (Table 5), soil texture fractions, vegetation parameters (Table 6) and meteorological data. At some sites, such as Bondville and Santarem Km67/Km83, the global and local values for LAI and V_{cmax} were markedly different (Fig. 6), though for the majority of sites, global and local LAI values were quite close (Fig. 6a), whereas global V_{cmax} values were underestimated compared to local values (below dashed line in Fig. 6b). Overall, the MODIS LAI values were closer to the local values and in general, lower than global values (Fig. 6a).

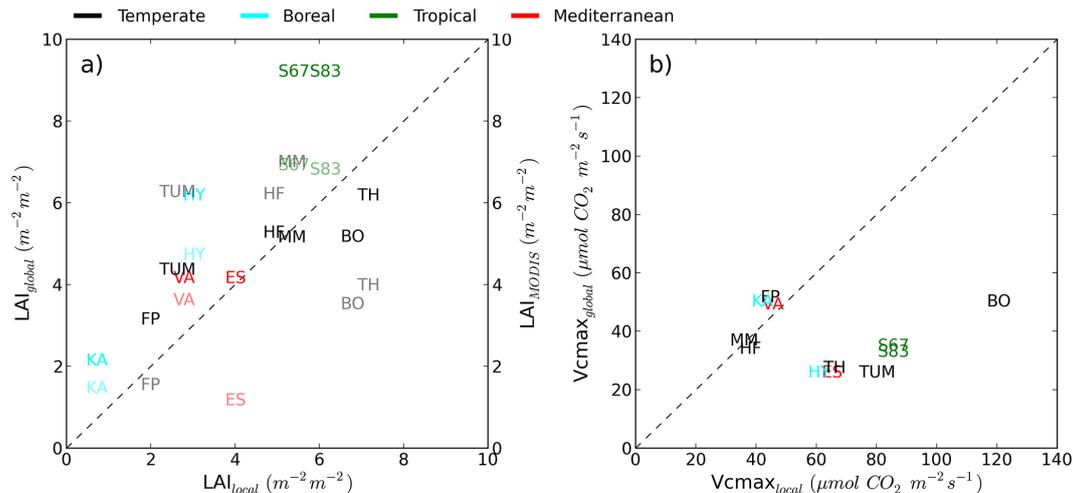


Figure 6. Comparison of (a) global, MODIS (site annual maximum) and local leaf area index (LAI) and (b) global and local maximum rate of Rubisco carboxylase activity (V_{cmax}) at the 12 FLUXNET sites (HF: Harvard Forest; VA: Vaira Ranch; MM: Morgan Monroe; HY: Hyytiälä; TH: Tharandt; TUM: Tumbarumba; ES: El Saler; FP: Fort Peck; KA: Kaamanen; S67: Santarem Km67; S83: Santarem Km83; BO: Bondville). The LAI data displayed for each study site refer to the annual maximum LAI of the dominant PFT. The site labels are coloured according to their climate zone (Table 2) and in (a), the lighter shades are the MODIS data. The dashed grey lines represent LAI and V_{cmax} , where global, MODIS and local values match, with overestimated global and MODIS values above the dashed line and underestimated values below it.

In general, we found the meteorological data to play a more important role than model parameters in determining GPP fluxes at sites, such as Santarem Km67 and Santarem Km83. At these sites, the meteorological forcing data was the primary driver of productivity and biases associated with the global meteorological data compensated for incorrect parameter values. However, at Tumbarumba, incorrectly predicted GPP was due to model error rather than meteorological data or model parameters. We performed a temperature sensitivity study at Tumbarumba using local meteorological and parameter data sets (local-F; Table 3). The winter and spring surface air temperatures (May–October) of the FLUXNET data were increased by increments of 1°C and the model was re-ran each time. Improvements in simulated seasonal cycle were observed, but only at high surface air temperatures (an increase in 7°C). Since the model performed *poorly* when using both global and local data meteorological data, we can assume that this is due to the model itself rather than the forcing data. Tumbarumba is classified as a sclerophyll forest and JULES does not have this land cover type. We assigned the Needleleaf (NL) PFT to JULES at this site. The introduction of the correct PFT and associated parameters may improve the results at this site.

4.2 Of the global meteorological data sets used in this study which one compares best to FLUXNET data?

At the majority of sites, the WFDEI data set compares better to local meteorological measurements (FLUXNET) than the PRINCETON data set does (Fig. 4). This is likely due

to the WFDEI data set being derived from the ECMWF Re-analysis (ERA-Interim) data set (Dee et al., 2011). The ERA-Interim re-analysis is a higher resolution data set ($\sim 0.75^{\circ} \times 0.75^{\circ}$; equivalent to a surface resolution of about $83\text{ km} \times 83\text{ km}$ at the Equator and $83\text{ km} \times 48\text{ km}$ at 55°N) than the NCEP-NCAR re-analysis ($2.0^{\circ} \times 2.0^{\circ}$; equivalent to a surface resolution of about $222\text{ km} \times 222\text{ km}$ at the Equator and $222\text{ km} \times 128\text{ km}$ at 55°N), from which the PRINCETON data set is derived (Kistler et al., 2001). The ERA-Interim re-analysis also uses a more advanced data assimilation system than the NCEP-NCAR re-analysis (Kistler et al., 2001; Weedon et al., 2014).

At the sites considered, differences between global and local values for downward short-wave and long-wave radiation fluxes and surface air temperatures are quite small (Fig. 4a, b and d), with average percentage RMSEs ranging from 0.4–17% (expressed as percentages of the annual mean value), while larger differences are observed for precipitation (Fig. 4c), with average percentage RMSEs ranging from 112–178%. At the majority of sites, there is a negative bias associated with precipitation (Fig. 4c), but this will have little effect on GPP fluxes since JULES is relatively insensitive to precipitation (Galbraith et al., 2010). For the remaining meteorological variables, there is a positive surface air temperature bias, but no dominant bias associated with the radiation fluxes. However, at individual sites, such as the tropical sites, Santarem Km67 and Santarem Km83, biases in the meteorological data can affect model results.

4.3 Are improvements in simulated GPP observed when forcing JULES with daily satellite phenology compared to using the default phenology module?

In general, we found that using MODIS data resulted in only small decreases in RMSE at a limited number of sites, compared to using locally observed LAI. At sites where model performance improved, improvements were a result of setting the annual maximum LAI to be the annual maximum MODIS LAI rather than forcing the model with daily MODIS LAI. The largest improvements in simulated GPP occur at sites with low annual LAI, such as the grassland (Vaira Ranch, Fort Peck, Kaamanen) and cropland (Bondville) sites and the tropical sites (Santarem Km67 and Santarem Km83). At the boreal sites, Tharandt and Hyytiala, the MODIS LAI tended to be quite noisy and this led to underestimated GPP (Fig. 5c).

We found that at sites where the MODIS LAI time-series was noisy (large day-to-day variations), this resulted in decreased model performance. At some of the flux tower sites, the MODIS data failed to capture aspects of the seasonal cycle of leaf phenology, such as the magnitude of the seasonal cycle (Tharandt, El Saler) and the beginning and end of the growing season (Bondville). For example, at Tumarumba, the MODIS instrument estimated the annual maximum LAI to be $6.08 \text{ m}^{-2} \text{ m}^{-2}$ and the daily LAI to be quite noisy, whereas the ground level observations show it to be $2.5 \text{ m}^{-2} \text{ m}^{-2}$ (Table 6) and LAI to be constant for much of the year.

The MODIS instrument provides a valuable source of information that can be used by land surface models. However, in this study, the quality of the LAI data can affect model performance. At the tropical sites, MODIS was unable to capture the magnitude of seasonal variation in LAI with MODIS overestimating the locally observed annual maximum LAI at Santarem Km67 and Santarem Km83 by 28 and 10%, respectively (Table 6). It was also unable to correctly capture LAI during the Amazonian rainy season, which runs from December to June, as a result of increased cloud cover. The MODIS LAI is very noisy in these regions, but should be constant throughout the year.

Overall, we found JULES' phenology module performed *very well* at the temperate sites and *poorly* at the tropical and cropland sites. The ability of the phenology model to simulate GPP fluxes *very well* at temperate sites, with slight underestimation of the summer carbon uptake and phase shift (leaf onset and senescence), may be due to its design; temperature-dependent for the BL/NL PFT classes, with model parameters tuned for temperate regions. Forcing the model with MODIS LAI only slightly improved model performance. However, setting the annual maximum LAI for each PFT to be the annual maximum MODIS LAI resulted in improved model performance, without the computational overhead of forcing JULES with daily satellite data. More accurate GPP predictions could be possible with the inclusion

of tropical PFTs, such as tropical evergreen broadleaf and tropical deciduous broadleaf, with associated model parameters and a phenology model modified to take these tropical PFTs into account.

5 Conclusions

We performed a multi-site evaluation of the JULES LSM using local, global and satellite data. In general, we found that when using local meteorological and parameter data sets, JULES performed *very well* (Fig. 2a and Table 4) at temperate sites with a negative bias observed at the tropical and cropland sites. At a limited number of sites, the model was able to simulate interannual variability using local data, with the exception of the tropical site, Santarem Km67, and Tumarumba.

The use of global data worsens model performance by introducing negative biases into model simulations of GPP at the majority of sites, with the exception of the tropical sites. The improvement in model simulated GPP when using local values of vegetation properties implies that global values may be incorrect. At sites where model performance improved using global data, this was due to biases associated with the meteorological data. We observed that the meteorological data had a greater impact on modelled GPP fluxes than model parameters.

The use of meteorological data extracted from global meteorological data sets was used to drive JULES. We found that global meteorological data increased the negative biases of local model simulations at all sites with the exception of the tropical sites, where GPP predictions were improved. Of the two global meteorological data sets used in this study, the WFDEI data set more closely captures the local meteorological conditions, though we found that the PRINCETON data set results in improved performance at some of the sites due to positive biases associated with the downward radiation fluxes and surface air temperature. This implies that there are compensating errors within the model, which need to be identified and addressed.

LAI is an important parameter used in the calculation of canopy photosynthesis. Model simulations using local and MODIS data displayed improvements in modelled GPP compared to using only local data. Improvements in modelled GPP were observed at the beginning and ending of the growing season. Using MODIS data for the annual maximum LAI allows for improved model performance without the complication of assimilating daily satellite data into the model. We found the default phenology module allowed JULES to perform *very well* at temperate sites, but not at the tropical sites. More realistic simulation of the seasonal cycle of GPP was observed at sites with low LAI values, such as grasslands. Even though we have described the MODIS data as being noisy at a number of sites, it provides a valuable source of information as it is a high spatial and temporal resolution

data set. It allows a better understanding of plant response to climate and is a useful aid to modellers.

Although only a limited number of model parameters were modified at the 12 flux tower sites, due to limited data availability at FLUXNET sites, we showed that with more accurate information regarding flux tower sites, improved predictions of GPP are possible. However, negative biases still exist in this situation due to model error and incorrect

modelling of tropical processes. We suggest that improved model performance with regards to the terrestrial carbon cycle could be achieved with the introduction of more PFT classes, such as tropical evergreen broadleaf and tropical deciduous broadleaf, and their associated model parameters and a phenology model that can properly simulate carbon fluxes in both temperate and tropical regions.

Appendix A: Deriving global model parameters used by the global operational version of JULES

In the Global Land Cover Characterization (version 2) database (GLCC), land cover is classified into 17 categories using the International Geosphere–Biosphere Programme (IGBP) scheme. Each flux tower has a land cover category assigned to it in the GLCC database (IGBP code in Table 5). These IGBP codes are then used to derive the annual maximum LAI (Table A1) and canopy height factor (Table A2) for each PFT. The canopy height (metres) is calculated from the canopy height factor (metres) and annual maximum LAI by using Eq. (A1):

$$\text{Canopy height} = \text{Canopy height factor} \times \text{LAI}^{\frac{2}{3}}. \quad (\text{A1})$$

Table A1. Annual maximum leaf area index (LAI) of JULES vegetation land cover types (PFTs) (BL: broadleaf tree; NL: needleleaf tree; C3g: C3 grass; C4g: C4 grass; sh: shrubs) for each of the 17 IGBP categories. Note that for the snow and ice, barren and water bodies categories, there are no LAI values available.

IGBP code	IGBP class	Leaf area index of JULES PFTs				
		BL	NL	C3g	C4g	sh
1	EN forest		6.0	2.0		
2	EB forest	9.0		2.0	4.0	
3	DN forest		4.0	2.0		
4	DB forest	5.0		2.0	4.0	3.0
5	Mixed forest	5.0	6.0	2.0		
6	Closed shrub			2.0		3.0
7	Open shrub	5.0		2.0	4.0	2.0
8	Woody savannah	9.0		4.0		2.0
9	Savannah	9.0			4.0	
10	Grassland			3.0	4.0	3.0
11	Permanent wetland	9.0		3.0		3.0
12	Cropland	5.0		5.0	4.0	3.0
13	Urban					
14	Crop/natural mosaic	5.0	6.0	4.0	4.0	3.0
15	Snow and ice					
16	Barren					
17	Water bodies					

Table A2. Canopy height factor (metres) of JULES vegetation land cover types (PFTs) (BL: broadleaf tree, NL: needleleaf tree, C3g: C3 grass, C4g: C4 grass, sh: shrubs).

	BL	NL	C3g	C4g	sh
Canopy height factor	6.5	6.5	0.5	0.5	1.0

Acknowledgements. We thank the anonymous reviewers and editor of this paper whose thoughtful suggestions helped considerably improve and clarify this manuscript. This work was carried out during D. Slevin's PhD research and was funded by a school scholarship from the School of Geosciences, University of Edinburgh. For local parameter data, we would like to thank the following people: Danilo Dragoni (Morgan Monroe); Bill Munger (Harvard Forest); Thomas Grünwald (Tharandt); Pasi Kolari (Hyytiälä); Eva Van Gorsel (Tumbarumba); and Ray Leuning (Tumbarumba). Chris Jones and Andy Wiltshire at the UK Met Office kindly provided information on the global parameter data used by the global operational version of JULES. Graham Weedon (UK Met Office) provided advice on using the WFDEI data set. The authors gratefully acknowledge advice from Tristan Quaipe (Reading University) and Jose Gómez-Dans (University College London) on using the MODIS data and Doug Clark at the Centre for Ecology and Hydrology (CEH) on model spin-up. We thank Carley Iles for reading the manuscript and providing feedback.

Edited by: G. Folberth

References

- Abramowitz, G.: Towards a public, standardized, diagnostic benchmarking system for land surface models, *Geosci. Model Dev.*, 5, 819–827, doi:10.5194/gmd-5-819-2012, 2012.
- Abramowitz, G., Leuning, R., Clark, M., and Pitman, A.: Evaluating the performance of land surface models, *J. Climate*, 21, 5468–5481, 2008.
- Alton, P., Mercado, L., and North, P.: A sensitivity analysis of the land-surface scheme JULES conducted for three forest biomes: biophysical parameters, model processes, and meteorological driving data, *Global Biogeochem. Cycles*, 20, GB1008, doi:10.1029/2005GB002653, 2007.
- Ashworth, K., Wild, O., and Hewitt, C. N.: Sensitivity of isoprene emissions estimated using MEGAN to the time resolution of input climate data, *Atmos. Chem. Phys.*, 10, 1193–1201, doi:10.5194/acp-10-1193-2010, 2010.
- Aurela, M., Tuovinen, J.-P., and Laurila, T.: Carbon dioxide exchange in a subarctic peatland ecosystem in northern Europe measured by the eddy covariance technique, *J. Geophys. Res.-Atmos.*, 103, 11289–11301, 1998.
- Baldocchi, D., Falge, E., Gu, L. H., Olson, R., and Hollinger, D.: 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.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., and Leroy, M.: LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION, Part 1: Principles of the algorithm, *Remote Sens. Environ.*, 110, 275–286, 2007.
- Beck, H. E., McVicar, T. R., van Dijk, A. I. J. M., Schellekens, J., de Jeu, R. A. M., and Bruijnzeel, L. A.: Global evaluation of four AVHRR–NDVI data sets: Intercomparison and assessment against Landsat imagery, *Remote Sens. Environ.*, 115, 2547–2563, doi:10.1016/j.rse.2011.05.012, 2011.
- Beerling, D. J. and Quick, W. P.: A new technique for estimating rates of carboxylation and electron transport in leaves of C3 plants for use in dynamic global vegetation models, *Glob. Change Biol.*, 1, 289–294, 1995.
- 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., Gash, J., Lloyd, A., Pryor, M., Weedon, G. P., and Shuttleworth, J.: Evaluating the JULES land surface model energy fluxes using FLUXNET data, *J. Hydrometeorol.*, 11, 509–519, 2010.
- Blyth, E., Clark, D. B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M., and Sitch, S.: A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, *Geosci. Model Dev.*, 4, 255–269, doi:10.5194/gmd-4-255-2011, 2011.
- Bonan, G. B.: Forests and climate change: forcings, feedbacks, and the climate benefits of forests, *Science*, 320, 1444–1449, 2008.
- Bruno, R. D., Da Rocha, H. R., De Freitas, H. C., Goulden, M. L., and Miller, S. D.: Soil moisture dynamics in an eastern Amazonian tropical forest, *Hydrol. Processes*, 20, 2477–2489, 2006.
- Cadule, P., Friedlingstein, P., Bopp, L., Sitch, S., Jones, C. D., Ciais, P., Piao, S. L., and Peylin, P.: Benchmarking coupled climate–carbon models against long-term atmospheric CO₂ measurements, *Global Biogeochem. Cycles*, 24, GB2016, doi:10.1029/2009GB003556, 2010.
- Cao, M. and Woodward, F. I.: Dynamic responses of terrestrial ecosystem carbon cycling to global climate change, *Nature*, 393, 249–252, 1998.
- Carson, D. J.: Current parameterizations of land-surface processes in atmospheric general circulation models, in: *Land Surface Processes in Atmospheric General Circulation Models*, Cambridge University Press, Cambridge, UK, 67–108, 1982.
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, in: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., pp. 465–570, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.
- 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.
- Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W.: Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sens. Environ.*, 106, 285–304, 2007.
- Cox, P. M.: Description of the TRIFFID Dynamic Global Vegetation Model, Hadley Centre Technical Note, 24, 1–16, 2001.

- 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. 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.
- Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.: The Common Land Model, *B. Am. Meteorol. Soc.*, 84, 1013–1023, 2003.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, 2011.
- Deng, F., Chen, M., Plummer, S., and Pisek, J.: Algorithm for Global Leaf Area Index Retrieval Using Satellite Imagery, *IEEE T. Geosci. Remote Sens.*, 44, 2219–2229, doi:10.1109/TGRS.2006.872100, 2006.
- Dickinson, R. E.: Biosphere/Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model, Technical report, 1986.
- Diner, D. J., Martonchik, J. V., Borel, C., Gerstl, S. A. W., Gordon, H. R., Myneni, R., Pinty, B., and Verstraete, M. M.: MISR Level 2 Surface Retrieval Algorithm Theoretical Basis, Jet Propulsion Laboratory, California Institute of Technology, Pasadena., pp. 1–81, 2008.
- Dirmeyer, P. A., Xiang, G., Zhao, M., Zhichang, G., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel analysis and implications for our perception of the land surface, *B. Am. Meteorol. Soc.*, 87, 1381–1397, 2006.
- Domingues, T. F., Martinelli, L. A., and Ehleringer, J. R.: Ecophysiological traits of plant functional groups in forest and pasture ecosystems from eastern Amazonia, Brazil, *Plant Ecol.*, 193, 101–112, 2007.
- Doughty, C. E. and Goulden, M. L.: Seasonal patterns of tropical forest leaf area index and CO₂ exchange, *J. Geophys. Res.-Biogeo.*, 113, G00B06, doi:10.1029/2007JG000590, 2008.
- El Maayar, M., Price, D. T., Black, T. A., Humphreys, E. R., and Jork, E.-M.: Sensitivity tests of the integrated biosphere simulator to soil and vegetation characteristics in a Pacific coastal coniferous forest, *Atmos. Ocean*, 40, 313–332, 2002.
- Farquhar, G. D., Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO₂ assimilation in leaves of C3 species, *Planta*, 149, 78–90, 1980.
- Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S. D., Tucker, C., Scholes, R. J., Le, Q. B., Bondeau, A., Eastman, R., Epstein, H., Gaughan, A. E., Hellden, U., Mbow, C., Olsson, L., Paruelo, J., Schweitzer, C., Seaquist, J., and Wessels, K.: Greenness in semi-arid areas across the globe 1981–2007 – An Earth Observing Satellite based analysis of trends and drivers, *Remote Sens. Environ.*, 121, 144–158, doi:10.1016/j.rse.2012.01.017, 2012.
- Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate–carbon cycle feedback analysis: results from the C4MIP model intercomparison, *J. Climate*, 19, 3337–3353, 2006.
- Galbraith, D., Levy, P. E., Sitch, S., Huntingford, C., Cox, P., Williams, M., and Meir, P.: Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change, *New Phytol.*, 187, 647–665, 2010.
- Gao, F., Morissette, J. T., Wolfe, R. E., Ederer, G., Pedelty, J., Masuoka, E., Myneni, R., Tan, B., and Nightingale, J.: An algorithm to produce temporally and spatially continuous MODIS-LAI time series, *IEEE Geosci. Remote S.*, 5, 60–64, 2008.
- Gilmanov, T. G., Tieszen, L. L., Wylie, B. K., Flanagan, L. B., Frank, A. B., Haferkamp, M. R., Meyers, T. P., and Morgan, J. A.: Integration of CO₂ flux and remotely-sensed data for primary production and ecosystem respiration analyses in the Northern Great Plains: potential for quantitative spatial extrapolation, *Global Ecol. Biogeogr.*, 14, 271–292, 2005.
- Gornall, J. L., Jónsdóttir, I. S., Woodin, S. J., and Van der Wal, R.: Arctic mosses govern below-ground environment and ecosystem processes, *Oecologia*, 153, 931–941, 2007.
- Goulden, M. L., Miller, S. D., Da Rocha, H. R., Menton, M. C., de Freitas, H. C., e Silva Figueira, A. M., and de Sousa, C. A. D.: Diel and seasonal patterns of tropical forest CO₂ exchange, *Ecol. Appl.*, 14, 42–54, 2004.
- Grünwald, T. and Bernhofer, C.: A decade of carbon, water and energy flux measurements of an old spruce forest at the Anchor Station Tharandt, *Tellus B*, 59, 387–396, 2007.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181–3210, doi:10.5194/acp-6-3181-2006, 2006.
- Hagen, S. C., Braswell, B. H., Linder, E., Frolking, S., Richardson, A. D., and Hollinger, D. Y.: Statistical uncertainty of eddy flux–based estimates of gross ecosystem carbon exchange at Howland Forest, Maine, *J. Geophys. Res.-Atmos.*, 111, 1984–2012, 2006.
- Hallgren, W. and Pitman, A.: The uncertainty in simulations by a Global Biome Model (BIOME3) to alternative parameter values, *Glob. Change Biol.*, 6, 483–495, 2000.
- Haverd, V., Leuning, R., Griffith, D., van Gorsel, E., and Cuntz, M.: The turbulent Lagrangian time scale in forest canopies constrained by fluxes, concentrations and source distributions, *Bound.-Lay. Meteorol.*, 130, 209–228, 2009.
- Henderson-Sellers, A., McGuffie, K., and Pitman, A. J.: The Project for Intercomparison of Land-Surface Parametrization Schemes (PILPS): 1992 to 1995, *Clim. Dynam.*, 12, 849–859, 1996.
- Hu, J., Su, Y., Tan, B., Huang, D., Yang, W., Schull, M., Bull, M. A., Martonchik, J. V., Diner, D. J., Knyazikhin, Y., and Myneni, R. B.: Analysis of the MISR LAI/FPAR product for spatial and temporal coverage, accuracy and consistency, *Remote Sens. Environ.*, 107, 334–347, 2007.

- Hutyra, L. R., Munger, J. W., Saleska, S. R., Gottlieb, E., Daube, B. C., Dunn, A. L., Amaral, D. F., De Camargo, P. B., and Wofsy, S. C.: Seasonal controls on the exchange of carbon and water in an Amazonian rain forest, *J. Geophys. Res.-Biogeo.*, 112, G03008, doi:10.1029/2006JG000365, 2007.
- Ito, A., Sillman, S., and Penner, J. E.: Global chemical transport model study of ozone response to changes in chemical kinetics and biogenic volatile organic compounds emissions due to increasing temperatures: Sensitivities to isoprene nitrate chemistry and grid resolution, *J. Geophys. Res.-Atmos.*, 114, 1984–2012, D09301, doi:10.1029/2008JD011254, 2009.
- Jones, C. P.: Unified Model Documentation Paper No 70: Ancillary File Data Sources, Tech. rep., UK Met Office, 2007.
- Kattge, J., Knorr, W., Raddatz, T., and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, *Glob. Change Biol.*, 15, 976–991, 2009.
- Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jeanne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation, *B. Am. Meteorol. Soc.*, 82, 247–267, 2001.
- Knorr, W.: Annual and interannual CO₂ exchanges of the terrestrial biosphere: process-based simulations and uncertainties, *Global Ecol. Biogeogr.*, 9, 225–252, 2000.
- Knorr, W. and Heimann, M.: Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme, *Global Biogeochem. Cy.*, 15, 207–225, 2001.
- Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneeth, A., Barr, A., Stoy, P., and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation, *Glob. Change Biol.*, 16, 187–208, 2010.
- Laurila, T., Soegaard, H., Lloyd, C. R., Aurela, M., Tuovinen, J.-P., and Nordstroem, C.: Seasonal variations of net CO₂ exchange in European Arctic ecosystems, *Theor. Appl. Climatol.*, 70, 183–201, 2001.
- Lawrence, P. J. and Chase, T. N.: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), *J. Geophys. Res.-Biogeo.*, 112, G01023, doi:10.1029/2006JG000168, 2007.
- Leakey, A. D. B., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P., and Ort, D. R.: Elevated CO₂ effects on plant carbon, nitrogen, and water relations: six important lessons from FACE, *J. Experimental Botany*, 60, 2859–2876, doi:10.1093/jxb/erp096, 2009.
- Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Mankut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster, U., Sitch, S., Takahashi, T., Viovy, N., van der Werf, G. R., and Woodward, F. I.: Trends in the sources and sinks of carbon dioxide, *Nat. Geosci.*, 2, 831–836, 2009.
- Leuning, R., Cleugh, H. A., Zegelin, S. J., and Hughes, D.: Carbon and water fluxes over a temperate Eucalyptus forest and a tropical wet/dry savanna in Australia: measurements and comparison with MODIS remote sensing estimates, *Agr. Forest Meteorol.*, 129, 151–173, 2005.
- Liu, Q., Gu, L., Dickinson, R. E., Tian, Y., Zhou, L., and Post, W. M.: Assimilation of satellite reflectance data into a dynamical leaf model to infer seasonally varying leaf areas for climate and carbon models, *J. Geophys. Res.-Atmos.*, 113, D19113, doi:10.1029/2007JD009645, 2008.
- Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *Int. J. Remote Sens.*, 21, 1303–1330, 2000.
- Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., Niu, S. L., Norby, R., Piao, S. L., Qi, X., Peylin, P., Prentice, I. C., Riley, W., Reichstein, M., Schwalm, C., Wang, Y. P., Xia, J. Y., Zaehle, S., and Zhou, X. H.: A framework for benchmarking land models, *Biogeosciences*, 9, 3857–3874, doi:10.5194/bg-9-3857-2012, 2012.
- Masson, V., Champeaux, J.-L., Chauvin, F., Meriguet, C., and Lacaze, R.: A Global Database of Land Surface Parameters at 1-km Resolution in Meteorological and Climate Models, *J. Climate*, 16, 1261–1282, 2003.
- Meyers, T. P. and Hollinger, S. E.: An assessment of storage terms in the surface energy balance of maize and soybean, *Agr. Forest Meteorol.*, 125, 105–115, 2004.
- Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., Batjes, N., Dijkshoorn, K., van Engelen, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E., and Shi, X.: Harmonized World Soil Database v1.2, Tech. rep., International Institute for Applied Systems Analysis (IIASA), Food and Agriculture Organization of the United Nations (FAO), 2012.
- Norby, R. J., DeLucia, E. H., Gielen, B., Calfapietra, C., Giardina, C. P., King, J. S., Ledford, J., McCarthy, H. R., Moore, D. J. P., Ceulemans, R., De Angelis, P., Finzi, A. C., Karnosky, D. F., Kubiske, M. E., Lukac, M., Pregitzer, K. S., Scarascia-Mugnozza, G. E., Schlesinger, W. H., and Oren, R.: Forest response to elevated CO₂ is conserved across a broad range of productivity, *P. Natl. Acad. Sci. USA*, 102, 18052–18056, doi:10.1073/pnas.0509478102, 2005.
- Piao, S., Friedlingstein, P., Ciais, P., Zhou, L., and Chen, A.: Effect of climate and CO₂ changes on the greening of the Northern Hemisphere over the past two decades, *Geophys. Res. Lett.*, 33, L23402, doi:10.1029/2006GL028205, 2006.
- Pitman, A. J.: The evolution of, and revolution in, land surface schemes designed for climate models, *Int. J. Climatol.*, 23, 479–510, 2003.
- Pugh, T. A. M., Ashworth, K., Wild, O., and Hewitt, C. N.: Effects of the spatial resolution of climate data on estimates of biogenic isoprene emissions, *Atmos. Environ.*, 70, 1–6, doi:10.1016/j.atmosenv.2013.01.001, 2013.
- Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowald, N. M., Lindsay, K., Lee, Y.-H., Nevison, C. D., Doney, S. C., Bonan, G., Stöckli, R., Covey, C., Running, S. W., and Fung, I. Y.: Systematic assessment of terrestrial biogeochemistry in coupled climate–carbon models, *Glob. Change Biol.*, 15, 2462–2484, 2009.

- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havráňková, 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, doi:10.1111/j.1365-2486.2005.001002.x, 2005.
- Ryu, Y., Baldocchi, D. D., Ma, S., and Hehn, T.: Interannual variability of evapotranspiration and energy exchange over an annual grassland in California, *J. Geophys. Res.*, 113, D09104, doi:10.1029/2007JD009263, 2008.
- Schaefer, K., Schwalm, C. R., Williams, C., Arain, M. A., Barr, A., Chen, J. M., Davis, K. J., Dimitrov, D., Hilton, T. W., Hollinger, D. Y., Humphreys, E., Poulter, B., Raczka, B. M., Richardson, A. D., Sahoo, A., Thornton, P., Vargas, R., Verbeeck, H., Anderson, R., Baker, I., Black, T. A., Bolstad, P., Chen, J., Curtis, P. S., Desai, A. R., Dietze, M., Dragoni, D., Gough, C., Grant, R. F., Gu, L., Jain, A., Kucharik, C., Law, B., Liu, S., Lokipitiya, E., Margolis, H. A., Matamala, R., McCaughey, J. H., Monson, R., Munger, J. W., Oechel, W., Peng, C., Price, D. T., Ricciuto, Dan Riley, W. J., Roulet, N., Tian, H., Tonitto, C., Torn, M., Weng, E., and Zhou, X.: A model–data comparison of gross primary productivity: results from the North American Carbon Program site synthesis, *J. Geophys. Res.-Biogeo.*, 117, G03010, doi:10.1029/2012JG001960, 2012.
- Schimel, D.: Carbon cycle conundrums, *P. Natl. Acad. Sci. USA*, 104, 18353–18354, 2007.
- Schmid, H. P., Grimmer, C. S. B., Cropley, F., Offerle, B., and Su, H.-B.: Measurements of CO₂ and energy fluxes over a mixed hardwood forest in the mid-western United States, *Agr. Forest Meteorol.*, 103, 357–374, 2000.
- Sellers, P. J., Mintz, Y., Sud, Y. C., and Dalcher, A.: A Simple Biosphere Model (SiB) for Use within General Circulation Models, *J. Atmos. Sci.*, 43, 505–531, 1986.
- Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *J. Climate*, 19, 3088–3111, 2006.
- Stöckli, R., Lawrence, D. M., Niu, G. Y., Oleson, K. W., Thornton, P. E., Yang, Z. L., Bonan, G. B., Denning, A. S., and Running, S. W.: Use of FLUXNET in the Community Land Model development, *J. Geophys. Res.*, 113, G01025, doi:10.1029/2007JG000562, 2008.
- Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Kulmala, M., and Vesala, T.: Long-term measurements of surface fluxes above a Scots pine forest in Hyytiälä, southern Finland, 1996–2001, *Boreal Environ. Res.*, 8, 287–302, 2003.
- Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO₂ exchange on timescales from hourly to decadal at Harvard Forest, *J. Geophys. Res.*, 112, G02020, doi:10.1029/2006JG000293, 2007.
- Warren, J. M., Pötzelsberger, E., Wullschleger, S. D., Thornton, P. E., Hasenauer, H., and Norby, R. J.: Ecohydrologic impact of reduced stomatal conductance in forests exposed to elevated CO₂, *Ecohydrology*, 4, 196–210, doi:10.1002/eco.173, 2011.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydrometeorol.*, 12, 823–848, 2011.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resour. Res.*, 50, 7505–7514, doi:10.1002/2014WR015638, 2014.
- 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.
- Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N. V., Wang, Y., Privette, J. L., Huemmrich, K. F., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D. E., Gower, S. T., Nemani, R. R., Knyazikhin, Y., and Myneni, R. B.: MODIS leaf area index products: from validation to algorithm improvement, *IEEE T. Geosci. Remote*, 44, 1885–1898, 2006.