Improvement of modelling plant responses to low soil moisture in JULESvn4.9 and evaluation against flux tower measurements

Anna B. Harper¹,², Karina E. Williams³,⁴, Patrick C. McGuire⁵,⁶, Maria Carolina Duran Rojas¹, Debbie Hemming³,⁷, Anne Verhoef⁶, Chris Huntingford⁸, Lucy Rowland⁴, Toby Matthews⁸, Cléon Breder Eller⁴, Camilla Mathison³, Rodolfo L.B. Nóbrega⁹, Nicola Gedney¹⁰, Pier Luigi Vidale⁵, Fred Otul-Larbi¹¹, Divya Pandey¹², Sebastien Garrigues¹³, Azin Wright⁶, Darren Slevin¹⁴, Martin G. De Kauwe¹⁵,¹⁶,¹⁷, Eleanor Blyth⁸, Jonas Aröd¹⁸, Andrew Black¹⁹, Damien Bonal¹⁹, Nina Buchmann¹⁰, Benoit Burban¹¹,²¹,²², Kathrin Fuchs²³, Agnès de Grandcourt²⁴,²⁵, Ivan Mammarella²⁶, Lutz Merbold²⁷, Leonardo Montagnani²⁸, Yann Nouvellon²⁴,²⁵, Natalia Restrepo-Coupe²⁹,³⁰, Georg Wohlfahrt³¹

¹ College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter, UK
² Global Systems Institute, University of Exeter, Exeter, UK
³ UK Met Office, Fitzroy Road, Exeter, UK
⁴ College of Life and Environmental Sciences, University of Exeter, Exeter, UK
⁵ Department of Meteorology and National Centre for Atmospheric Science, University of Reading, Reading RG 66 BB, UK
⁶ Department of Geography & Environmental Science, University of Reading, Reading RG 66 BB, UK
⁷ Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK
⁸ Centre for Ecology and Hydrology, Wallingford, UK
⁹ Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, Berkshire, SL5 7PY, UK
¹⁰ Met Office Hadley Centre, Joint Centre for Hydrometeorological Research, Maclean Building, Wallingford OX10 8BB, United Kingdom
¹¹ Lancaster Environment Centre, Lancaster University, LA1 4YQ, UK
¹² Stockholm Environment Institute at York, University of York, York, UK
¹³ ECMWF, Copernicus Atmospheric Monitoring Service, Reading, UK
¹⁴ Forest Research, Northern Research Station, Roslin, Midlothian, EH25 9SY, UK
¹⁵ ARC Centre of Excellence for Climate Extremes, Sydney, NSW 2052, Australia.
¹⁶ Climate Change Research Centre, University of New South Wales, Sydney, NSW 2052, Australia.
¹⁷ Evolution & Ecology Research Centre, University of New South Wales, Sydney, NSW 2052, Australia.
¹⁸ Department of Physical Geography and Ecosystem Science, Lund University Sölvegatan 12 S-223 62 Lund Sweden
¹⁹ Université de Lorraine, AgroParisTech, INRAE, UMR Silva, 54000 Nancy, France
²⁰ Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland
²¹ INRAE, AgroParisTech, CIRAD,
²² CNRS, Université de Guyane, Université des Antilles, UMR Ecologique Fonctionnelle et Biogéochimie des Sols & Agro-écosystèmes, F-34060 Montpellier, France
²³ Institute of Meteorology and Climate Research - Atmospheric Environmental Research, Karlsruhe Institute of Technology (KIT Campus Alpin), 82467 Garmisch-Partenkirchen, Germany
²⁴ CRDPI, BP 1291 Pointe Noire, Republic of Congo
²⁵ CIRAD, UMR Eco&Sol, Ecologie Fonctionnelle & Biogéochimie des Sols & Agro-écosystèmes, F-34060 Montpellier, France
²⁶ Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Finland
²⁷ International Livestock Research Institute (ILRI), Mazingira Centre, PO Box 30709, 00100 Nairobi, Kenya
²⁸ Faculty of Science and Technology, Free University of Bolzano, Bolzano, Italy; Forest Services, Autonomous Province of Bolzano, Bolzano, Italy
²⁹ Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, 85721 USA
³⁰ School of Life Science, University of Technology Sydney, Sydney, NSW, 2006 Australia
Abstract

Drought is predicted to increase in the future due to climate change, bringing with it a myriad of impacts on ecosystems. Plants respond to drier soils by reducing stomatal conductance, in order to conserve water and avoid hydraulic damage. Despite the importance of plant drought responses for the global carbon cycle and local/regional climate feedbacks, land surface models are unable to capture observed plant responses to soil moisture stress. We assessed the impact of soil moisture stress on simulated gross primary productivity (GPP) and latent energy flux (LE) in the Joint UK Land Environment Simulator (JULES) vn4.9 on seasonal and annual timescales, and evaluated ten different representations of stress in the model. For the default configuration, GPP was more realistic in temperate biome sites than in the tropics or high latitudes/cold region sites, while LE was best simulated in temperate and high latitude/cold sites. Errors not due to soil moisture stress, possibly linked to phenology, contributed to model biases for GPP in tropical savannah and deciduous forest sites. We found that three alternative approaches to calculating soil moisture stress produced more realistic results than the default parameterization for most biomes and climates. All of these involved increasing the number of soil layers from 4 to 14, and the soil depth from 3m to 10.8m. In addition, we found improvements when soil matric potential replaced volumetric water content in the stress equation, when the onset of stress was delayed, and when roots extended deeper into the soil. For LE, the biases were highest in the default configuration in temperate mixed forests, with overestimation occurring during most of the year. At these sites, reducing soil moisture stress (with the new parameterizations mentioned above) increased LE and made the simulation worse. Further evaluation into the reason for the high bias in LE at many of the sites would enable improvements in both carbon and energy fluxes with new parameterizations for soil moisture stress.

1 Introduction

Drought has a range of impacts on terrestrial ecosystems (Allen et al., 2010; Choat et al., 2012), plays a role in feedbacks on the weather and climate systems across scales (Seneviratne et al., 2013; Lemordant et al., 2016; Miralles et al., 2019; Lian et al., 2020) and affects the global carbon cycle (Green et al., 2017; Humphrey et al., 2018; Peters et al., 2018). These impacts and feedbacks have the potential to affect society, either directly through moisture availability effects on crops, or indirectly by adjusting near-surface temperatures, or forcing large-scale variations to the climate system. Roughly 40% of the vegetated land surface is limited by seasonal water deficits (Nemani et al., 2003; Beer et al., 2010), which are a major control on gross primary productivity (GPP) in sub-humid, semi-arid, and arid regions (Stocker et al., 2018). In the future, soil moisture stress for ecosystems is predicted to increase over large regions (Berg et al., 2016; Ukkola et al., 2020). In this paper, we define
“soil moisture stress” to mean the physiological stress experienced by vegetation due to its interactions with unusually dry soils. Feedbacks from water-limited ecosystems on the global carbon cycle may make it more difficult to stabilize the climate system at global warming thresholds such as two degrees. For these reasons, accurate process-based models of plant response to soil moisture stress are needed in coupled land-atmosphere climate models. However, the models used to represent biogeophysical and biogeochemical processes in Earth System Models (ESMs) are often unable to properly capture observed responses to soil moisture stress (e.g. Beer et al., 2010; Powell et al., 2013; Medlyn et al., 2016; Restrepo-Coupe et al., 2017; De Kauwe et al., 2017; Peters et al., 2018; Paschalis et al., 2020).

Plants respond to reductions in soil moisture content (SMC) through a range of drought tolerance and prevention strategies. Commonly, plants respond to low SMC by reducing their stomatal aperture to conserve water and protect the xylem from damage (Field and Holbrook, 1989; Sparks and Black, 1999). Cavitation and embolism usually happened when the plant is under water stress and the water potential in the xylem is low enough to fill the xylem with water vapor and/or air instead of water (Lambers et al., 2008; Choat et al., 2012). Embolized xylem is unable to transport water, and for some vegetation types, this is a dominant cause of plant mortality under drought conditions (e.g. (Brodrribb and Cochard, 2009). To avoid this, many plants limit water loss by reducing their stomatal conductance when soil moisture levels reach a certain threshold (Tyree and Sperry, 1989; Sperry et al., 1998; Choat et al., 2012) or by shedding leaves (Wolfe et al., 2016). High atmospheric vapor pressure deficits (VPD), which sometimes occur in conjunction with meteorological drought, may also result in stomatal closure. The reduced stomatal conductance triggers a cascade of other responses, beginning with reduced rates of photosynthesis (Ball et al., 1987), which reduce carbon uptake and possibly growth (Merbold et al., 2009a; Doughty et al., 2015). The lower stomatal conductance will reduce transpiration, which causes more surface available energy to be converted into sensible heat. This transference of latent to sensible heat can contribute to increased land surface temperature and amplification of heat waves (Seneviratne et al., 2010). Over the long term, droughts can lead to changes in plant species composition (Liu et al., 2018) or large-scale forest mortality (McDowell et al., 2008), sometimes causing a transient situation where large ecosystems switch from being a sink of carbon dioxide to a source (Ciais et al., 2005; Gatti et al., 2015).

There is a spectrum of mechanisms through which species tolerate or acclimate to drought, meaning a “one-size-fits-all” approach to modelling can be inadequate. Explicit model representations of the xylem hydraulics are complex and difficult to parameterize globally. The emergence of plant trait databases has enabled early models to represent the hydraulic properties of the soil-plant-atmosphere continuum (e.g. (Sperry et al., 2016; Eller et al., 2018; De Kauwe et al., 2020; Eller et al., 2020; Sabot et al., 2020). Also, new approaches are emerging that focus on ‘plant profit maximisation’, where photosynthetic uptake of CO₂ is optimally traded against plant hydraulic function, as an alternative to the empirical functions commonly used in models to regulate gas exchange during periods of water stress (Sabot et al., 2020).

More often, for now, land surface models (LSMs) represent the regulation of stomatal conductance as a simple generic function of SMC, generally expressed in terms of volumetric water content (θ, m³ m⁻³). This simple generic function is the so-called “beta” function, where β is a factor between zero and one that limits photosynthesis in some way (depending on the model, See Methods). Above a critical SMC, there is no stress (β=1), and below the critical point, stress increases as SMC
decreases, until the wilting point is reached ($\beta=0$). Alternative, yet related, expressions are available whereby stomatal regulation occurs through changes in the soil matric potential ($\psi$, expressed in pressure units, such as MPa); $\theta$ and matric potential (a measure of “how tightly the water is held in the soil pores”, thereby affecting rootwater uptake) are closely related via the water retention curve. However, using one function for all plant responses to drying soils can result in errors, and especially in the extreme cases, for example, by underestimating stress in wet ecosystems and overestimating stress in dry ecosystems. In fact, the parameters describing hydraulic response to soil moisture may change in time (Robinson et al., 2019), and can vary between ecosystem types (Teuling et al., 2010). Such variation may be in response to climate change, or evolving vegetation and soil properties, and their structure.

In this study, we focus on the effects of droughts on vegetation that occur due to low SMC. Although droughts are often associated with changes beyond low precipitation levels, including high air temperatures and VPD, these climate drivers have their own set of impacts on vegetation, adding to the effects of low SMC, that will not be addressed here. We explore different ways in which soil moisture stress can be represented in a widely used model of the terrestrial biosphere, the Joint UK Land Environment Simulator (JULES) (Best et al., 2011; Clark et al., 2011). JULES is a community model, and is used in coupled or standalone mode, forced by meteorological variables. Its applications are on timescales ranging from weather forecasting to climate projections, and the model is the terrestrial component of the UK Earth System Model and the HadGEM family of models (Martin et al., 2011). The spatial scales are similarly diverse. Studies range from single-point modelling of crop yield at one site (Williams et al., 2017), which requires detailed knowledge of one crop variety under carefully controlled conditions, to global predictions of land sources and sinks of CO$_2$ for the annually updated Global Carbon Project (Friedlingstein et al., 2019), which requires reliable performance for all vegetation types across the globe. Soil moisture stress has been identified as a key driver of variability in JULES projections (Blyth et al., 2011). (Verhoef and Egea, 2014) showed that the standard $\beta$ function in JULES, and similar LSMs, needs urgent attention, as to whether it is the most appropriate functional form, and/or if parameterized correctly. In an evaluation of the model across ten flux tower sites, (Blyth et al., 2011) showed that the “dry-down” of the sites in semi-arid areas was too quick and the seasonal variation of evaporation in the tropics was too great, possibly due to the roots being modelled as too shallow (Blyth et al., 2011). This was later supported by studies that found that the root depths of LSMs were too shallow (Teuling et al., 2006; Wang and Dickinson, 2012). Indeed, some LSMs (CLM, SiB3, TERRA-ML) were able to improve model performance by representing deeper (e.g. 10m) and more efficient roots (Baker et al., 2008; Akkermans et al., 2012; Liu et al., 2020).

Evaluating the impact of simulated soil moisture stress on vegetation requires that other model errors that also affect CO$_2$ and water fluxes are minimized. For instance, it is possible that the rapid drying found in (Blyth et al., 2011) was due to over-estimation of soil evaporation. The fact that land surface models in general over-estimate evapotranspiration during wet periods is well documented (Blyth et al., 2011; Martínez-de la Torre et al., 2019) and leads to unrealistically low soil moisture after long dry periods (Ukkola et al., 2016). The high evaporation (and subsequent low SMC) could be due to errors in factors not being studied in this study, such as radiation absorption or turbulent exchanges with the atmosphere. Leaf area
index (LAI) also strongly affects the magnitude and seasonality of fluxes coming from vegetation and soil (via variations in shading).

This study aimed to evaluate the simulation of GPP and LE for a range of biomes and climates, to diagnose sites and seasons when soil moisture stress affects the results, and to evaluate different methods for representing soil moisture stress in JULES. To do this, we chose a subset of sites in the FLUXNET2015 database and from the Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) experiment based on availability of data. Where possible we prescribed soil moisture and LAI from site measurements, to differentiate the roles of SMC, the $\beta$ parameterization, or modelled phenology in model biases. We used the GPP calculated before soil moisture stress is applied to understand seasons and locations where the $\beta$ parameterization was contributing to model errors. We also reviewed other commonly used approaches for modelling soil moisture stress, presented in Section 2.2, to motivate the representations evaluated in the remainder of the paper. This work is one of the first published results from a JULES community-wide focus group (called a JULES Process Evaluation Group, or JPEG) on understanding soil moisture stress impacts on vegetation, which began in 2016.

2 Methods

2.1 Photosynthesis and stomatal conductance in JULES

The Joint UK Land Environment Simulator (JULES) (Best et al., 2011; Clark et al., 2011) is a process-based model that simulates the fluxes of carbon, water, energy and momentum between the land surface and the atmosphere. JULES treats each vegetation type as existing on a separate tile within a grid box. Energy and carbon flux calculations are performed separately for each tile, depending on Plant Functional Type (PFT)-dependent parameters. Leaf-level net photosynthesis ($A$) is integrated over the canopy, according to the canopy radiation scheme specified. In the present study, we use 10 canopy layers of equal LAI (in JULES this is ‘canopy radiation model 6’), although another option in JULES is to use a ‘big leaf’ approach (Clark et al., 2011). Potential (non-stressed) photosynthesis is calculated based on three limiting rates: $W_c$ (a RuBisCO limited rate), $W_l$ (a light-limited rate), and $W_e$ (a transport limited rate for C3 plants and a PEPCarboxylase limitation for C4 plants). For full details on the photosynthesis scheme in JULES see (Clark et al., 2011; Harper et al., 2016).

Stomatal conductance to water vapour $g_s$ (in m s$^{-1}$) is related to net photosynthesis $A$ (in molCO$_2$ m$^{-2}$ s$^{-1}$) through:

$$g_s = \frac{-1.6A}{c_i - c_a}$$

where $c_a$ and $c_i$ are the atmospheric and intercellular CO$_2$ concentrations, respectively, in Pa, and 1.6 is the molar diffusivity ratio of CO$_2$ to H$_2$O in air (e.g. Guerrieri et al., 2019). $R$ is the universal gas constant (8.314 J K$^{-1}$ mol$^{-1}$) and $T^*$ is the leaf temperature (K). Vapour deficit at the leaf surface ($D$) affects stomatal conductance through the gradient between $c_a$ and $c_i$:

$$\frac{c_i - c_a}{c_i - c^*} = f_0 \left(1 - \frac{D}{D_{crit}}\right)$$

Here, $c^*$ is the photorespiration compensation point (Pa), $D$ is the humidity deficit at the leaf surface (g kg$^{-1}$), and $D_{crit}$ and $f_0$ are PFT-dependent parameters (Cox et al., 1998; Best et al., 2011).
2.2 Soil moisture stress in JULES and other terrestrial biosphere models

Many land surface, terrestrial biosphere, and crop models include a $\beta$ function to represent the effect of soil moisture stress on vegetation. The stress factor can be applied in various places, with implementations generally split into two categories: stomatal and biochemical limitation (Bonan et al., 2014; De Kauwe et al., 2015). JULES falls under the latter category, with potential leaf-level carbon assimilation, $A_p$, being converted to the water-limited net leaf photosynthesis through multiplication with the stress factor:

$$ A = A_p \beta $$

Other land surface models apply biochemical limitation through reducing RuBisCO or reducing electron transport (e.g. ORCHIDEE, (Krinner et al., 2005)). CABLE applies limits to both the stomata (via reducing $g_s$) and $A$ (De Kauwe et al., 2015).

In JULES, soil moisture stress ($\beta$, unitless) for each soil layer $k$ is a function of volumetric water content ($\theta$) in each layer ($\theta_k, m^3 m^{-3}$) using:

$$ \beta_k = \begin{cases} \frac{1}{\theta_k - \theta_{wilt,k}} & \theta_k \geq \theta_{upp,k} \\ \frac{\theta_{wilt,k} - \theta_k}{\theta_{upp,k} - \theta_{wilt,k}} & \theta_{wilt,k} \leq \theta_k \leq \theta_{upp,k} \\ 0 & \theta_k \leq \theta_{wilt,k} \end{cases}, $$

where $\theta_{wilt}$ and $\theta_{upp}$ are the water contents at the wilting point and at which the plant starts to become water stressed, respectively (Cox et al., 1998). $\theta_{upp}$ is a function of $\theta_{crit}$, the critical water content (usually defined as the field capacity), and $p_0$, a PFT-dependent parameter:

$$ \theta_{upp} = \theta_{wilt} + (\theta_{crit} - \theta_{wilt})(1 - p_0) $$

The parameter $p_0$ was recently added to JULES (in version 4.6) to allow $\beta$=1 for $\theta<\theta_{crit}$, in other words delaying the onset of stress as soils initially dry below the field capacity. In the default configuration, $p_0$ is set to 0 (meaning $\theta_{upp} = \theta_{crit}$), and $\theta_{wilt}$ and $\theta_{crit}$ correspond to matric soil water potentials of -1.5 MPa and -0.033 MPa, respectively. Equation 4 means that, for each soil layer, soil moisture stress completely limits root water extraction from that layer if $\theta_k$ is at or below the wilting point ($\beta_k=0$), and there is no soil moisture stress ($\beta_k=1$) if $\theta_k$ is at or above $\theta_{upp,k}$. In between these points, there is a linear increase in stress (decrease in $\beta_k$) as water content decreases (blue line in Fig. 1). The overall soil moisture stress factor is calculated based on the water in each soil layer and the fraction of root mass in that soil layer $r_k$ (as the latter co-determines root water extraction, see Eq. 7):

$$ \beta = \sum_{k}^{n\text{soil}} r_k \beta_k $$

JULES has four soil layers ($n_{soil} = 4$) that together extend to 3 meters depth. The root mass distribution is assumed to decay exponentially with depth $z$, i.e. $e^{-z/d}$ where $d_i$ is a PFT-specific parameter that gives the e-folding depth of root (Fig. 2). The stress factor is also applied to leaf maintenance respiration (and optionally to stem and root maintenance respiration), and affects the fraction of total plant transpiration extracted from each soil layer, $\epsilon_k$: 
\[ \varepsilon_k^0 = \frac{r_k \beta_k}{\beta} \]  

2.2 Alternative representations of soil moisture stress

2.2.1 Deeper soil column and roots (soil14 and soil14_deeproots)

Several studies have found that deep roots are an essential part of modelling plant drought responses (e.g. (Teuling et al., 2006; Baker et al., 2008; Akkermans et al., 2012; Wang and Dickinson, 2012). In the ‘soil14’ experiments, \( n_{\text{soil}} \) increased from 4 to 14, and the thickness of each soil layer \( dz_{\text{soil}} \) was changed as in Table 1, to give a total depth of 10.8m. The parameter \( d_r \) remained unchanged, giving the root profiles shown in the colored lines in the middle panel of Fig. 2. In ‘soil14_deeproots’, we doubled \( d_r \), which gave more emphasis to deeper layers in Eqs. 6 and 7.

2.2.2 Delayed onset of stress (p0 and soil14_p0)

Measurements of transpiration rates show that plants do not limit transpiration until intermediate levels of soil dryness occur (Verhoef and Egea, 2014). In JULES, a delay of stress can be represented with the parameter \( p_0 \) (Eq. 5). In the ‘p0’ experiments, we used \( p_0 = 0.4 \) to capture the delayed onset. This was done with both the 4 layer (p0) and 14 layer (soil14_p0) soils.

2.2.3 Curvilinear response (psi and soil14_psi)

While Eq. 4 assumes a linear increase in stress as water content decreases, some models assume a curvilinear increase in stress (Sinclair, 2005; Oleson et al., 2010; Egea et al., 2011), or an S-shaped curve (Tardieu and Davies, 1993; De Kauwe et al., 2015). Nonlinear responses can be represented by a parameter to induce curvature (Egea et al., 2011) or through using the soil matric potential rather than \( \theta \):

\[ \beta_{\psi_k} = \frac{\psi_k - \psi_{\text{close}}}{\psi_{\text{open}} - \psi_{\text{close}}} \]  

Models that use soil water potential include (Verhoef and Egea, 2014; Fatichi et al., 2012; Manzoni et al., 2013; Lawrence et al., 2019), while other models use leaf water potential (e.g. (Tuzet et al., 2003; Christina et al., 2017)). In the ‘psi’ experiments, we replaced Eq. 4 with Eq. 8, and set \( \psi_{\text{open}} \) and \( \psi_{\text{close}} \) to -0.033 and -1.5 MPa, respectively. This was done with both the 4 layer (psi) and 14 layer (soil14_psi) soils.

2.2.4 Remove root-weighted access to soil moisture (mod1 and soil14_mod1)

Consideration of root distribution and thickness of soil layers, and related depth of soil profile, also varies across models. The measure of water availability for \( \beta \) can be a function of each layer’s water content (Eq. 6), or water in the wettest layer (Martens et al., 2017), or the contribution of the water in each layer can be weighted by the root density or plant and soil hydraulics (Oleson et al., 2010; Christina et al., 2017). Another approach is to use a function of water in the whole root
column (\(\bar{\theta}\)), rather than layer-by-layer, which is equivalent to assuming that plants can access water anywhere in the soil column, as long as there are roots present (Baker et al., 2008; Harper et al., 2013):

\[
\beta_{\text{mod1}} = \frac{\bar{\theta} - \bar{\theta}_{\text{wilt}}}{\theta_{\text{upp}} - \theta_{\text{wilt}}} \tag{9}
\]

In this approach, root water extraction per layer is weighted by layer thickness (\(dZ_{\text{soil}}\)) rather than by beta:

\[
e_{k}^{\beta} = dZ_{\text{soil}}(\theta_{k} - \theta_{\text{wilt}}) \tag{10}
\]

In the ‘mod1’ experiments, Eqs. 4 and 6 were replaced with Eq. 9; and Eq. 7 was replaced with Eq. 10. In addition, \(d_{r}\) was implemented as the maximum root depth instead of the e-folding depth and was double its default value (with a maximum depth of 3m). Root fraction in each soil layer was set equal to the layer thickness. In ‘soil14_mod1’, \(d_{r}\) was double its default value (Table 2), but without enforcing a maximum depth of 3m, resulting in root profile given by the dashed line in Fig. 2 (middle panel).

### 2.2.5 Exponential decline of roots with depth (soil14_realroots)

The root profile from grasses with \(n_{\text{soil}}=14\) and depth of 10.8m more closely resembles the often observed rapid decay of root biomass with depth than the profiles for other PFTs. We evaluated the impact of using more realistic root distributions by setting \(d_{r}\) to 0.5 for all PFTs in the ‘soil14_realroots’ experiment. Essentially, this gave more emphasis to shallow layers in calculating root water extraction and \(\beta\), and is an opposite approach of the ‘mod1’ experiments, which gave more emphasis to the thickest soil layers.

### 2.2.6 Stomatal limitation

Although not used in this study, it is worth noting that many land surface and terrestrial biosphere models apply soil moisture stress through limiting stomatal conductance (Egea et al., 2011; Fatichi et al., 2012; De Kauwe et al., 2015). These include JSBACH and DLEM ((Radatz et al., 2007; Tian et al., 2010). For example, CABLE uses \(\beta\) to modify the slope of the relationship between stomatal conductance and net photosynthesis (De Kauwe et al., 2015). In other models (e.g. crop model WOFOST), they interact through allowing the actual or potential evapotranspiration to impact the soil moisture threshold for unstressed vegetation (Tardieu and Davies, 1993). Models with stomatal limitations from soil moisture can include the explicit consideration of the plant/soil hydraulics (Williams et al., 1996; Zhou et al., 2013; Bonan et al., 2014; Mirfenderesgi et al., 2016; Eller et al., 2018; Kennedy et al., 2019; De Kauwe et al., 2020) and/or chemical signalling, such as the abscisic acid (ABA) concentration in the xylem sap (Tardieu and Davies, 1993; Dewar, 2002; Verhoef and Egea, 2014; Huntingford et al., 2015; Takahashi et al., 2018). In other models, \(\beta\) can affect root growth and leaf senescence (e.g. (Arora and Boer, 2005; Song et al., 2013; Wang et al., 2016)), or reduce mesophyll conductance (Keenan et al., 2010).

### 2.3 Model set up and evaluation

We evaluated JULES at 42 sites covering eight general biome types from the tropics to the Arctic (SM Table 1). Each JULES simulation was run with meteorological measurements taken at each site (i.e. point-scale runs rather than simulating the entire gridbox). The meteorological and flux tower observations were obtained from the LBA Model Intercomparison.
Project (sites with 'LBA' in the name) or FLUXNET2015 dataset (Pastorello et al., 2020). At each site, we extracted temperature, precipitation, wind speed, surface pressure, specific humidity, longwave and shortwave radiation for running JULES at either half-hourly or hourly resolution, depending on the data available. We then used measured LE and calculated GPP as supplied in both datasets. The FLUXNET2015 data was filtered for periods with low u* based on the variable Ustar threshold (VUT) method. NEE was partitioned into GPP and Reco based on the nighttime method of Reichstein et al. (2005).

All fluxes and meteorological variables were previously gap-filled following the MDS method in Reichstein et al. (2005). The LBA data was also gap-filled and u* filtered using site-specific u* thresholds (Restrepo-Coupe et al., 2013). In the FLUXNET2015 data, a relationship between Reco and Ta was parameterized based on nighttime data and applied to the whole dataset to partition NEE into GPP and Reco (Reichstein et al., 2005), while in the LBA dataset Reco was derived based on the u* filtered nighttime NEE (not including temperature which was not observed to correlate with NEE at night, for details see (Restrepo-Coupe et al., 2013).

For sites with soil moisture data, we individually contacted site PIs to gather details on the depth of the measurements and other details on soil texture, physical properties, and root depth. This resulted in a subset of 21 sites with soil moisture measurements. We also collected data from site PIs on LAI. Fourteen sites had both LAI and SM data. Often the time period of LAI/SM measurements was shorter than the full record, and we only ran JULES for the time periods with the most data to avoid the need for gap-filling. The time periods of the simulations and soil layers for prescribing data are provided in Table SM1.

The default plant parameter set was taken from (Harper et al., 2016). When LAI was not prescribed, we used the JULES phenology scheme to predict LAI. This scheme predicts leaf growth and senescence based on temperature alone. Tile fractions were determined from the vegetation class (Table SM1). We calculated soil properties from information supplied by site PIs where possible, otherwise we used the gridbox sand, silt, clay fractions of the Met Office Central Ancillary Program (CAP) high resolution input file (Dharssi et al. 2009) to derive the Brooks and Corey (1964) parameters, along with the approximations of the parameters (via pedotransfer function) required for the soil hydraulic properties as detailed in Cosby et al. (1984). The soil properties are available with the files used to run JULES (see data availability). Each simulation began with a 50-year spin-up of the soil moisture using recycled meteorology.

This evaluation focused on seasonal and annual timescales of fluxes. We started with daily measurements from the sites, then masked any modelled outputs on days when measurements were not available and calculated monthly means when >50% of the data was present. To evaluate the model performance, we used four metrics: normalized absolute error (NAE), variance ratio (VR), correlation coefficient (r), and root mean error (RME). The NAE compares flux estimates to measured values averaged over the entire simulation period. The other metrics are calculated from monthly mean fluxes. The VR is the ratio of variance in the simulations to the observations. For a perfect fit, the VR would be 1: lower values mean the model variance is too low, and vice versa (Carvalhais et al., 2008). R is the Pearson’s correlation coefficient and it gives an indication of model-data agreement on both a seasonal and year-to-year timescale. For the soil moisture stress experiments, we use Taylor diagrams and RMSE based on annual mean fluxes to evaluate the best fit.
3. Results

3.1 Simulated GPP and ET

On average, JULES matched the pattern of observed seasonal cycle well for sites in non-agricultural biomes in temperate and cold climates (mean $r>0.80$). The correlation was fairly good (mean $r>0.60$) for sites in tropical grasslands and savannas, and cropland. However, the seasonal cycle was not well represented for sites in tropical dry forests (mean $r=0.43$) or tropical evergreen forests (mean $r = -0.10$).

In terms of model biases, the NAE was lowest (mean <0.2) for GPP at tropical evergreen forest, cropland, temperate grassland, and woody savannah sites, while NAE was highest in tropical grassland and cold grassland sites (mean >0.50). The variance ratio (VR) indicates the amount of simulated variability in comparison to observations. On average, VR was between 0.5–1.5 for sites in tropical savannahs, temperate non-agricultural biomes, and boreal forest. A low VR indicates that simulated variability was too low – this was the case for sites in cold grasslands and cropland. Conversely, a high VR indicates that simulated variability was higher than observed, and was seen for sites in tropical forests and grasslands. In this case, the error was due to a higher than observed seasonal cycle (ie LBA-K67 in Fig. 5).

Certain patterns emerged with regards to the skill of the JULES model to predict observed GPP within the biomes which occur in multiple climates (evergreen needle-leaf forests, grasslands, and savannahs), indicating that the model tends to perform best in temperate midlatitude climates. First, the bias was higher and the correlation was lower at the sites with a Mediterranean climate (e.g. IT-CA1) compared to the other temperate mixed forest sites. Second, sites in temperate grasslands were better simulated than those in cold and tropical grasslands. The NAE values were more than three and six times as high for sites in cold and tropical grasslands, respectively, compared to those in temperate grasslands. Third, NAE was significantly higher for sites in tropical savannahs compared to those in more temperate savannah sites in the US.

The model performance was also more related to climate than biome for LE, with the largest biases occurring for sites in the tropics. On average, the seasonal cycle of LE was well simulated for sites outside of the tropics (mean $r$ per biome > 0.8), and for sites in tropical savannahs ($r=0.79$). The NAE was lowest for sites in temperate grasslands and croplands (<0.05), followed by tropical evergreen and dry forest sites, and tropical savannah sites. Model variance was close to observed for the tropical savannah sites, but it was too high for the other tropical biome sites.

3.2 Role of soil moisture stress in GPP errors

At the sites with the poorest model performance for GPP (evergreen broadleaf sites, Mediterranean climates, cold and tropical grasslands, and tropical savannahs), we compared the GPP that JULES would calculate if there was no soil moisture stress to the actual simulated GPP (Fig. 5, Fig SM5). This was possible through a new diagnostic added to the model, which output GPP prior to multiplication by $\beta$. At the tropical forest sites (GF-Guy, LBA-K34, LBA-K67, LBA-K83, and LBA-
BAN), simulated GPP (standard approach) decreased during the dry season, while the unstressed GPP and observed GPP remained high during dry seasons, which indicates that the model was over-estimating soil moisture stress during the dry season. At the tropical grassland and savannah sites (AU-Fog, LBA-PDG, LBA-K77, and LBA-FNS), the modelled GPP (standard approach) was often too high, and the unstressed GPP was even higher. An exception was ZA-Kru, where the observed GPP was somewhere in between simulated GPP and unstressed GPP. There were mixed results for the sites with a Mediterranean climate (IT-CA1, US-Ton, US-Var); stress was impacting the GPP but other processes were also affecting the simulation. For example, at IT-CA1 the modelled GPP was very close to measured values when observed soil moisture and LAI were used, indicating that errors in soil hydrology and phenology were important at this site. At other semi-arid sites (IT-Col, US-Ton, US-Var), the bias occurred during the peak growing season, when JULES GPP was lower than observed.

In the cold grassland sites, soil moisture stress sometimes resulted in too low GPP (e.g. RU-Che), potentially because plants could not access frozen soil moisture. Other sites where JULES showed a large improvement with the unstressed GPP were the Aspen site in Canada (CA-Oas), Tharandt evergreen needleleaf forest in Germany (DE-Tha), the deciduous broadleaf forest in Belgium (BE-Vie), and the cropland site (US-Ne1). This analysis gives a list of sites that are useful for further exploring the role of soil moisture status in vegetation functioning: all sites with a Mediterranean climate or in tropical forests, as well as ZA-Kru, RU-Che, CA-OAs, DE-Tha, BE-Vie, and US-Ne1.

When prescribing soil moisture and LAI (see Sect. 2.3), the general trends in model performance were the same as before, although often the simulated GPP was less realistic with more prescribed data. This could be due to compensating errors within JULES (i.e. with regards to soil physical parameterizations related to infiltration or soil evaporation, see also (Van den Hoof et al., 2013)). The simulations at the tropical evergreen forest sites still did not resemble the measured GPP (as indicated by very low or negative correlations), even with prescribed LAI and soil moisture. It is possible that soil layers below those typically measured are influencing the forests soil water balance and canopy exchange processes, so more data are needed to accurately prescribe the full soil moisture profile. Only 14 sites had enough data to prescribe both soil moisture and LAI from site observations (Sect. 2.3), and often the time resolution of data was monthly which for soil moisture could miss impact of extremely wet or dry periods. However, most often adding the LAI data resulted in an improved simulation of GPP, indicating biases resulting from the JULES phenology scheme. The improvements with incorporation of prescribed LAI were particularly large for the cropland sites, and at LBA-RJA, which is a seasonally dry tropical forest.

We categorize the sites depending on the impact of soil moisture stress on their simulation of GPP with the most available prescribed data (for example, in the simulation with soil moisture and LAI prescribed at LBA-BAN, and for the simulation with soil moisture only at CN-HaM). The four categories are:

- Sites with underestimated GPP: Simulated GPP was lower than observed. However, $\beta$ was often 1, and removing soil moisture stress had a small effect on the simulation, indicating the importance of other processes in regulating GPP at these sites. Two tropical (LBA-K34, LBA-RJA) and two temperate grasslands (AT-Neu, CH-Cha) sites fall into this category.
• Sites with overestimated GPP: Simulated GPP was higher than observed, so removing soil moisture stress increased GPP and made the simulation worse. This category includes one tundra site (CN-HaM), a Mediterranean woodland (IT-CA1), two coniferous evergreen forests in Finland and Italy (FI-Hyy and IT-Ren), an arid grassland (US-SRG) and a tropical savannah (SD-Dem).

• Soil moisture stressed sites: As in the first set of sites, there was a low bias in GPP, but removing soil moisture stress improved the simulation. The “stressed” sites includes three temperate mixed forests (BE-Vie, DE-Tha, and US-UMB), a Mediterranean deciduous forest (IT-Col), a boreal Aspen forest (CA-Oas), a tropical evergreen forest (GF-Guy), and a cropland site (US-Ne1).

• Stressed sites plus other errors: At several sites, removing soil moisture stress made the simulation slightly better, but apparently other missing processes also affect the simulation. The difference between this category and the soil moisture stressed sites is the fact that there would still be a large bias even without soil moisture stress. Sites in this category include tropical forests (LBA-Ban, LBA-K83, LBA-K67), cropland (US-Ne2, US-Ne3), two savannah sites (ZA-Kru and US-SRM), and a tundra site (RU-Che).

The challenge is to determine a representation of soil moisture stress which improves the simulations at sites falling into categories 3 and 4 without degrading the simulation at the other sites. Clearly, we do not want to completely remove soil moisture stress as this plays an important role in regulating seasonal cycles in many ecosystems. In the remainder of the paper, we will focus on examples of changes at some of these sites.

3.3 New treatments of soil moisture stress

We ran the ten experiments (Section 2.2, Table 1) at a subset of sites that span the categories listed in Section 3.2. This included four sites where soil moisture stress leads to large model biases: GF-Guy, BE-Vie, DE-Tha, and CA-Oas. At these sites, there was an improvement when the 14 layer soil is combined with model settings p0, psi, or deeproots (representing, respectively, setting $p_0$ in Eq. 3 to 0.4; using Eq. 8, that depends on the soil matric potential to represent $\beta$; and having a deeper root profile). RMSE decreased from 2.23 gC m$^{-2}$ d$^{-1}$ on average to 1.47, 1.44, and 1.57 gC m$^{-2}$ d$^{-1}$, respectively in the soil14_p0, soil14_psi, and soil14_deeproots experiments, averaged across the four sites. There was also an improvement in the VR (from >2 for the default simulation to 0.92-1.0 for the above-listed experiments) and the correlation coefficient (from 0.74 to 0.86-0.88). For LE, the RMSE was slightly higher in these experiments (22.31, 22.22, and 20.41 W m$^{-2}$, respectively for soil14_p0, soil14_psi, and soil14_deeproots) compared to the default experiment (19.53 W m$^{-2}$), and the correlation coefficient was >0.82. Taylor diagrams for GPP and LE for all sites are shown in Figs. SM8 and SM11; and seasonal cycles of fluxes and simulated $\beta$ are shown in Figs. SM9, 10, and 11.

For example, at the tropical forest site (GF-Guy), experiments with default soil depths had correlation coefficients $r$<0.4, and an exaggerated seasonal cycle, as indicated by the high normalized standard deviation in the Taylor diagrams (Fig. 6). In the soil14_p0, soil14_psi, and soil14_deeproots experiments, the correlation $r$ was >0.7 (compared to 0.2 in the default
configuration), and the standard deviation was closer to observed. The GF-Guy site experienced the lowest amount of soil moisture stress in the soil14_p0 and soil14_psi experiments, which led to a more realistic simulation of GPP at this site (Fig. 7). Giving the model a more realistic root profile that decayed exponentially with depth (‘soil14_realroots’) produced the worst results. In the ‘soil14_realroots’ simulation, β was still weighted by root distribution, so the dry top soil layers had a relatively large impact on the stress experienced by the plants. This is further illustrated in Figure SM7 at the LBA-K67 site, which plots β against soil moisture in the top 1m. When \( d_r = 0.5 \text{m} \) (as in ‘soil14_realroots’), there were proportionally more roots in the top soil layers, and as these dried out, there was a sharp decline in β. In comparison, with \( d_r = 3 \text{m} \) (the default value) the trees were able to access water from deeper layers, so β did not decline as rapidly. As a result, β was very low during the dry season at the tropical forest sites in the ‘soil14_realroots’ experiments (Fig. 7). Another site in the ‘soil moisture stressed’ category is DE-Tha, where most simulations yield reasonable results (r>0.9). Only the default and ‘soil14_realroots’ simulations produced results outside the standard deviation of measured GPP. Variability (denoted by standard deviation in the Taylor diagram as well as VR close to 1) was best in the soil14_p0, p0, soil14_psi, and psi simulations.

At the sites with a Mediterranean climate (IT-Col, US-Var, US-Ton), soil14_psi and soil14_p0 removed the most stress, but p0 and psi with the default soil also produced a good fit. However the RMSE for LE was significantly higher in these experiments (22.55–26.09 W m\(^{-2}\) compared to 19.67 W m\(^{-2}\) in the default simulation), while the correlation coefficient was high for these four experiments (0.85–0.87 compared to 0.88 in the default).

At the sites with soil moisture stress plus other errors, there were fewer improvements although RMSE decreased from 2.81 gC m\(^{-2}\) d\(^{-1}\) in the default simulation to 2.08, 2.14, and 2.17 gC m\(^{-2}\) d\(^{-1}\) in the soil14_psi, soil14_p0, and soil14_deeproots simulations, respectively. These sites are LBA-K67, LBA-BAN, ZA-Kru, and RU-Che. The VR was best captured in the soil14_deeproots simulations, while the correlation coefficient was highest in the default simulation and in the soil14_realroots simulation. At LBA-K67, soil14_psi and soil14_p0 had the lowest RMSE and seasonal variation in GPP, although for all experiments the correlation coefficient was negative. At ZA-Kru, all results were within the range of the measurements, although the growing season GPP was underestimated. At LBA-BAN, soil14_deeproots, soil14_psi, and soil14_p0 gave lowest RMSE, but VR was very high (>3) and the correlation coefficient was low (r<0.4) for all simulations. There was very little difference between any of the simulations at RU-Che. For LE, there was a significant reduction in RMSE from 22.54 W m\(^{-2}\) to <18 W m\(^{-2}\) for all experiments with 14-layer soil (except for ‘soil14_realroots’) at these sites. The correlation coefficient was also significantly improved in these experiments (from 0.48 in the default simulation to >0.67).

Averaging across the ten sites where we performed additional experiments, the lowest RMSE occurred in the soil14_p0, soil14_psi, and soil14_deeproots experiments. The variability was best captured by the soil14, soil14_p0, and soil14_psi experiments (as denoted by VR close to 1). The mean correlation coefficient was similar across all experiments (0.50–0.56). For LE, averaged across all sites, the RMSE was similar across experiments, but VR was closest to 1 in the soil14 and
soil14_deeproots experiments. The correlation was highest (r~0.75–0.76 compared to default r=0.70) for all experiments with a 14-layer soil, except for the ‘reallroots’ experiment.

4. Discussion and Conclusions

Tables 3-4 summarize some of the key findings from this study. JULES simulated GPP was more realistic in temperate biome sites than in the tropics or high latitudes/cold region sites, as indicated by three statistics to measure annual biases (NAE), seasonal cycles (r), and variability (VR). LE was best simulated in temperate and high latitude/cold sites based on the same statistics (except for temperate mixed forests). For sites in the tropics, the default β parameterization contributed to an exaggerated seasonal cycle of GPP compared to the measurements, especially in tropical evergreen forests. Although the absolute error was low in tropical evergreen forest sites (LBA sites K34, K83, K67, and BAN for example), the seasonal cycle was overestimated (despite LAI being nearly constant all year), as indicated by high VR and low correlation coefficients. A similar result was observed with LE in most tropical sites: the seasonal cycle was incorrect and the VR was high. For example, at LBA-K67, the measurements show an increasing trend in GPP from August to October (coinciding with the dry season), while JULES predicted a decreasing trend during this time.

Even with soil moisture and LAI prescribed for the four tropical evergreen forest sites, the correlation coefficients were negative. At these sites, it is possible that including a seasonally varying photosynthetic capacity would improve the results, as in (Wu et al., 2017). The dry season is often accompanied by enhanced carbon uptake in Amazon forests, due to a combination of fewer clouds and increased incoming solar radiation (Saleska et al., 2003; Restrepo-Coupe et al., 2013; von Randow et al., 2013; Zeri et al., 2014) and seasonal leaf flushing (Wu et al., 2016). The observed seasonality in GPP is enabled by deep roots that can access ample soil moisture, and by the relatively high photosynthetic capacity of new leaves (Wu et al., 2017), a process not yet represented in JULES.

Other errors, possibly linked to phenology, also contributed to model biases in tropical savannah and deciduous forest sites. The improvements seen when LAI was prescribed at LBA-RJA (a seasonally dry tropical forest site) further suggest that JULES’ lack of a moisture-driven phenology scheme could be affecting the results at this site. LBA-RJA serves as interesting comparison to LBA-K67: RJA receives a similar amount of annual rainfall, but the dry season is more intense, with about half as much rainfall during the dry season compared to K67 (Restrepo-Coupe et al., 2013). The bedrock is relatively shallow at RJA (2–3 m, Christoffersen et al. (2014)), therefore deep soil moisture is not present. At this site, measured GPP drops steadily from January until reaching a minimum in the middle of the dry season. JULES captured this seasonal cycle very well, although the amplitude was slightly dampened with predicted GPP being higher than observed during most of the year (with prescribed LAI and soil moisture).

In cold grassland sites, JULES underpredicted the variability of GPP and had high annual biases. The biases were due to very little GPP being simulated, with β being low year-round. At RU-Che, giving more emphasis to deeper layers (in the ‘mod1’ experiment or with ‘soil14_deeproots’) did not increase GPP – which is not unexpected due to the presence of frozen
soils both in the simulations and in reality at this site (Merbold et al., 2009b). The C3 grass PFT at this site has most roots in the top 0.5m, which indicates that evaporation or sublimation could be drying the soils too much in the layers with the most roots and unfrozen soil moisture content.

We found that three alternative approaches to calculating soil moisture stress produced more realistic results than the default parameterization for most biomes and climates: 14-layer soil with a curvilinear stress response function (‘soil14_psi’, Eq. 8), 14-layer soil with delayed induction of stress (‘soil14_p0’, Eq. 3), and 14-layer soil with deeper roots (‘soil14_deeproots’). For LE, the biases were highest in the default configuration in temperate mixed forests, with overestimation occurring during Spring-Autumn. At these sites, reducing soil moisture stress (ie. with soil14_psi, soil14_p0, and soil14_deeproots) increased LE and made the simulation worse. Further evaluation into the reason for the high bias in LE at many of the sites would enable improvements in both carbon and energy fluxes with new parameterizations for $\beta$.

There is ample justification for having deeper soils and roots in JULES. Total soil column depth and root distribution determine the total amount of water and nutrients available to plants. Deep roots can access soil moisture at depth (Christina et al., 2017) and potentially the water table, and hence contribute to tree transpiration during dry periods, e.g. for GF-Guy where many canopy trees are not impacted by dry season droughts (Stahl et al., 2013b; Stahl et al., 2013a). Deep roots have been found to be important for many vegetation types and ecosystems (Canadell et al., 1996; Germon et al., 2020; Pierret et al., 2016): for multiple tree species in tropical forests (e.g. Nepstad et al., 1994; Jipp et al., 1998; Strey et al., 2017), for Acacias in semi-arid savannas such as SD-Dem (Ardö et al., 2008), and for fast-growing Eucalypt and Acacia mangium plantations in Brazil (Christina et al., 2011; Laclau et al., 2013; Germon et al., 2018), to name a few examples. These examples contrast with the shallow soils (3 meters) in the default JULES simulations. In addition, weighting root water uptake or soil moisture stress by fraction of roots in each layer could produce too much stress, if the shallow layers (with the most roots) dry out too quickly. Deep roots are very efficient at moving water, for example, specific hydraulic conductivities ($K_s$) of deep roots can be as much as 15 times higher than $K_s$ of superficial roots for Banksia sp (Pate et al., 1995), and deep roots can redistribute water from deep to shallow layers (Caldwell et al., 1998; Burgess et al., 2001; Oliveira et al., 2005). However, not all plants rely on deep roots during a drought (Prechsl et al., 2015; Brinkmann et al., 2019), and we did not find consistent improvements when weighting the contribution of each layer’s $\beta_i$ by the thickness of that layer (the ‘mod1’ experiments) rather than by the fraction of roots in that layer. This approach was similar to results from (Baker et al., 2008), where they found that increasing the soil column from 3.5m to 10m and allowing roots to access this entire reservoir improved the fit of the SiB3 model to observations at the LBA-K83 site. Similarly, the ability of the G’Day process-based model to accurately simulate wood production in fast-growing sub-tropical plantations was considerably improved by accounting for tree ability to uptake water in deep soil layers (Marsden et al., 2013). On the other hand, using the default calculation for $\beta$ with a more realistic root distribution parameterised by e-fold depth $d_e=0.5m$ increased overall soil moisture stress at most sites and produced a poor fit to measured GPP and LE.

In this study, we used flux tower observations and detailed site information when possible. Working with site researchers enabled us to narrow down reasons for model biases by prescribing soil moisture and LAI at some sites, and to better
understand mechanisms of drought responses at others. These are invaluable benefits of working with site-level data. However, there is potential to extract even more information from available datasets to improve the representation of soil moisture-vegetation interactions (Gentine et al., 2019). This includes better utilisation of satellite data, and one particular opportunity is to consider soil moisture measurements in parallel with those of solar-induced fluorescence, which is used to estimate photosynthesis (Lee et al., 2013). Satellite records have large spatial coverage, and modern machine learning algorithms could be used to characterise Earth Observed datasets of drought conditions (Huntingford et al., 2019). Such methods could address the difficulty in modelling the high complexity and geographical diversity of plant adaptive responses to soil moisture deficits that exist in nature.

Future work should build upon these results to further evaluate JULES response with these parameterizations, focusing on deeper soils and either using a non-zero $p_0$ (we used 0.4 in this study), or replacing Eqs. 4–5 with Eq. 8, (e.g., Wright et al in prep; Otu-Larbi et al in prep). We note that such alternative parameterizations are not a replacement for improved representations of the soil-plant hydraulic system, such as the SOX model that has recently been included in some versions of JULES (Eller et al. 2020).

The land is currently providing a partial offsetting of emissions by CO$_2$ photosynthetic drawdown, but this could be reversed if droughts increase in frequency or intensity in the future. Feedbacks from the land surface can amplify and lock-in existing drought conditions (Morillas et al., 2017), and land surface responses to regional drought can affect precipitation and circulation in other regions (Harper et al., 2013; Lian et al., 2020). Improving responses of vegetation to drought in land surface models such as JULES would have far-reaching implications for global climate modelling, and are therefore of utmost importance.

**Code Availability**

Both the model code and the files for running it are available from the Met Office Science Repository Service: [https://code.metoffice.gov.uk/](https://code.metoffice.gov.uk/). Registration is required and code is freely available subject to completion of a software licence. The results presented in this paper were obtained from running JULES branch [https://code.metoffice.gov.uk/trac/jules/browser/main/branches/dev/karinawilliams/r9227_add_gpp_unstressed_diagnostic](https://code.metoffice.gov.uk/trac/jules/browser/main/branches/dev/karinawilliams/r9227_add_gpp_unstressed_diagnostic), which is a branch of JULESv4.9 with the additional unstressed GPP diagnostic added. The runs were completed with the Rose suite [https://code.metoffice.gov.uk/trac/roses-u/browser/a/l/752/u-al752-jpegpaper](https://code.metoffice.gov.uk/trac/roses-u/browser/a/l/752/u-al752-jpegpaper), which also includes python scripts for creating the plots.

**Data Availability**

The FLUXNET2015 data used to run JULES is available for download from: [https://fluxnet.org/](https://fluxnet.org/).

**Author Contributions**
This study is the result of a large community effort within the JULES community to better understand soil moisture stress and simulated responses to soil moisture deficits. All co-authors contributed at some point to writing or improving the manuscript. Flux tower researchers provided particular insight into their sites: L. Montagnani (IT-Ren); I. M. (FI-Hyy); D. B. (GF-Guy); A. G. and Y. N. (CG-Tch); G. W. (AT-Neu); N. B., L. Merbold, K. F. (CH-Cha), L. Merbold (RU-Che, Za-Kru).

Competing Interests

The authors declare no competing interests.

Acknowledgements

Several site PIs contributed data for the soil moisture and LAI prescribed runs, we gratefully acknowledge their contribution here: Simone Sabbatini (IT-CA1), Sabina Keller (CH-Cha), Todd Schimelfening, David Scoby, and Tim Arkebauer (US-Ne1, US-Ne2, US-Ne3), Caroline Vincke (BE-Vie), Chris Gough (US-UMB), Tanguy Manise (BE-Vie), Pasi Kolari (FI-Hyy), Russ Scott (US-SRG and US-SRM, US-Whs, US-Wkg), and Alessandro Araujo (LBA). The JULES soil moisture stress group gratefully acknowledges Colin Prentice for input along the way. The authors also acknowledge the following funding: EPSRC Living with Environmental Change Fellowship EP/N030141/1 (A. B. H.); the Met Office Hadley Centre Climate Programme (HCCP) funded by BEIS and Defra (K. W., D. H., C. M.); NERC IMPETUS Project (NE/L010488/1) (A. V., A. W.); Newton Fund through the Met Office Climate Science for Service Partnership Brazil (CSSP Brazil) (K. W., N. G., C. M.; A. B. H., L. R.); the Research Endowment Trust Fund of the University of Reading (P. C. M.); National Aeronautics and Space Administration (NASA) LBA investigation CD-32, NASA LBA-DMIP project (# NNX09AL52G) (N. R.-C.); Province of South Tyrol “Cycling of carbon and water in mountain ecosystems under changing climate and land use (CYCLAMEN)” (G. W.); EU project SUPER-G (contract no. 774124) and the SNF project M4P (40FA40_154245) (N. B.); ERC under the EU’s Horizon 2020 research and innovation programme (grant agreement no. 787203 REALM) (R. N.); European project “Quantification, understanding and prediction of carbon cycle, and other GHG gases, in Sub-Saharan Africa” (CarboAfrica, STREP-CT-037132) (Y. N.); and the Newton/NERC/FAPESP Nordeste project: NE/N012488/1 (R. N., A. V.). The GF-Guy site is supported by an Investissement d’Avenir grant from the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-0025; ARBRE, ANR-11-LABX-0002-01). LBA data was provided with support from National Aeronautics and Space Administration (NASA) LBA investigation CD-32, NASA LBA-DMIP project (# NNX09AL52G), and the Gordon and Betty Moore Foundation “Simulations from the Interactions between Climate, Forests, and Land Use in the Amazon Basin: Modeling and Mitigating Large Scale Savannahization” project.
References


CRC Press


Table 1: Summary of the 10 JULES model experiments, related to the treatment of soil moisture stress.

<table>
<thead>
<tr>
<th>Experiment Name</th>
<th>Summary of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>Eq. 4 used for $\beta$. 4 layer soil to 3 m depth. Root profile in left panel of Fig. 2.</td>
</tr>
<tr>
<td>psi</td>
<td>Use matric soil potential (Eq. 8) rather than volumetric water content (Eq. 4) to calculate $\beta$; induces a curvilinear response.</td>
</tr>
<tr>
<td>p0</td>
<td>Reduce the critical VWC where stress begins. $p0$ in Eq. 5 is changed from 0 to 0.4 (green dashed line in Fig. 1).</td>
</tr>
<tr>
<td>mod1</td>
<td>Allow plants to access all soil moisture in the column. Eq. 9 replaces Eq. 4, and Eq. 10 replaces Eq. 7. Double default $d_r$ (max value 3). $d_r$ is the maximum depth of roots instead of e-folding depth.</td>
</tr>
<tr>
<td>soil14</td>
<td>Increase soil layers to 14, 10.8 m depth, but $d_r$ remains unchanged. Root profile in middle panel of Fig. 2.</td>
</tr>
<tr>
<td>soil14_deeproots</td>
<td>Increase soil layers to 14, 10.8 m depth, but double $d_r$ (gives more emphasis to deeper layers).</td>
</tr>
<tr>
<td>soil14_psi</td>
<td>Combine soil14 and psi experiments.</td>
</tr>
<tr>
<td>soil14_mod1</td>
<td>Combine soil14 and mod1 experiments, except $d_r$ is not capped at 3m. Root profile is the dashed line in middle panel of Fig. 2.</td>
</tr>
<tr>
<td>soil14_p0</td>
<td>Combine soil14 and p0 experiments.</td>
</tr>
<tr>
<td>soil14_realroots</td>
<td>Increase soil layers to 14, 10.8 m depth. Set $d_r=0.5$m for all PFTs, gives a more realistic reduction of root density with depth (see C3, C4 grass root profile in middle panel of Fig. 2).</td>
</tr>
</tbody>
</table>
Table 2. Default parameter settings (changed in experiments summarized in Table 1). In the JULES code, \( p_0 \) is called \texttt{fsmc\_p0}; \( n_{\text{soil}} \) is called \texttt{sm\_levels}; \( d_r \) is called \texttt{rootd}; \( \psi_{\text{open}} \) is \texttt{psi\_open}; \( \psi_{\text{close}} \) is \texttt{psi\_close}.

<table>
<thead>
<tr>
<th>JULES Parameter</th>
<th>Explanation</th>
<th>Default setting</th>
<th>Change in experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{fsmc_shape}</td>
<td>Switch that controls whether ( \beta ) decreases linearly with VWC ( \theta ) or with soil matric potential ( \psi ).</td>
<td>0</td>
<td>1 in psi and soil14_psi</td>
</tr>
<tr>
<td>\texttt{l_use_pft_psi}</td>
<td>Switch that controls whether ( \beta ) is a function of ( \theta_{\text{wilt}} ) and ( \theta_{\text{crit}} ) (false) or ( \psi_{\text{close}} ) and ( \psi_{\text{open}} ) (true)</td>
<td>false</td>
<td>true in psi and soil14_psi</td>
</tr>
<tr>
<td>( \psi_{\text{open}} )</td>
<td>Soil matric potential (MPa) above which ( \beta ) is 1. Dimension of \texttt{npft}.</td>
<td>None (only used when \texttt{l_use_pft_psi=true})</td>
<td>-0.033 for all PFTs in psi and soil14_psi</td>
</tr>
<tr>
<td>( \psi_{\text{close}} )</td>
<td>Soil matric potential (MPa) below which ( \beta ) is 0. Dimension of \texttt{npft}.</td>
<td>None (only used when \texttt{l_use_pft_psi=true})</td>
<td>-1.5 for all PFTs in psi and soil14_psi</td>
</tr>
<tr>
<td>( p_0 )</td>
<td>Threshold at which plants begin to feel stress (when \texttt{l_use_pft_psi=false}). Dimension of \texttt{npft}.</td>
<td>0</td>
<td>0.4 for all PFTs in ( p_0 ) and soil14_p0</td>
</tr>
<tr>
<td>\texttt{fsmc_mod}</td>
<td>Switch for method of weighting the contribution that each soil layer makes to the total ( \beta ). Dimension of \texttt{npft}.</td>
<td>0</td>
<td>1 for all PFTs in \texttt{mod1} and soil14_mod1</td>
</tr>
<tr>
<td>( d_r )</td>
<td>If \texttt{fsmc_mod}=0, ( d_r ) is the e-folding depth of roots assuming an exponential root distribution with depth. If \texttt{fsmc_mod}=1, ( d_r ) is the total depth of the root zone. Dimension of \texttt{npft}.</td>
<td>Tropical broadleaf evergreen trees = 3m Other broadleaf trees and deciduous needleleaf trees = 2m Evergreen needleleaf trees = 1.8m C3 and C4 grasses = 0.5m Shrubs = 1m</td>
<td>10.8 for all PFTs in \texttt{soil14_mod1} 0.5 for all PFTs in \texttt{soil14_realroots}</td>
</tr>
<tr>
<td>( n_{\text{soil}} )</td>
<td>Number of soil layers</td>
<td>4</td>
<td>14 in all soil14 experiments</td>
</tr>
<tr>
<td>( dz_{\text{soil}} )</td>
<td>Soil layer depths in meters, starting with the uppermost layer.</td>
<td>0.1, 0.25, 0.65, 2.0 (total depth = 3m)</td>
<td>0.1, 0.2, 0.2, 0.2, 0.3, 0.3, 0.3, 0.4, 0.4, 0.4,</td>
</tr>
<tr>
<td></td>
<td>1.0, 1.0, 3.0, 3.0 (total depth = 10.8m) in all soil14 experiments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.0, 1.0, 3.0, 3.0 (total depth = 10.8m) in all soil14 experiments
Table 3. Summary of model performance for GPP with no prescribed data. Ranking of seasonal cycle is based on the average Pearson’s correlation coefficient for each biome: Good (r≥0.80); Fair (0.80>r≥0.60); Poor (r < 0.60); Very poor (r≤0). Ranking of absolute error is based on the average normalized absolute annual error (NAE) for each biome: Low (NAE < 20%); Medium (NAE < 50%); High (NAE < 100%); and Very High (NAE > 100%).

<table>
<thead>
<tr>
<th>Climate</th>
<th>Biome</th>
<th>Seasonal Cycle</th>
<th>Absolute error (annual)</th>
<th>Variability</th>
<th>Diagnosed source of error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropics</td>
<td>Evergreen forests</td>
<td>Very poor</td>
<td>Low</td>
<td>Too high</td>
<td>Soil moisture stress during the dry season, or other phenological controls on GPP</td>
</tr>
<tr>
<td></td>
<td>Deciduous forests</td>
<td>Poor</td>
<td>Medium</td>
<td>Too high</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Grasslands</td>
<td>Fair</td>
<td>High</td>
<td>Too high</td>
<td>GPP is too high all year</td>
</tr>
<tr>
<td></td>
<td>Woody savannahs</td>
<td>Fair</td>
<td>Medium</td>
<td>Fair</td>
<td>--</td>
</tr>
<tr>
<td>Temperate</td>
<td>Mixed forests</td>
<td>Good</td>
<td>Medium</td>
<td>Fair</td>
<td>Soil moisture stress during growing season</td>
</tr>
<tr>
<td></td>
<td>Grasslands</td>
<td>Good</td>
<td>Low</td>
<td>Fair</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Woody savannahs</td>
<td>Good</td>
<td>Low</td>
<td>Fair</td>
<td>Multiple factors (soil moisture stress, hydrology, and phenology)</td>
</tr>
<tr>
<td></td>
<td>Cropland</td>
<td>Fair</td>
<td>Medium</td>
<td>Too low</td>
<td>--</td>
</tr>
<tr>
<td>High latitude or altitude</td>
<td>Boreal forests</td>
<td>Good</td>
<td>Medium</td>
<td>Fair</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Grasslands</td>
<td>Good</td>
<td>High</td>
<td>Too low</td>
<td>Frozen soils</td>
</tr>
</tbody>
</table>
Table 4. Summary of model performance for LE with no prescribed data. Ranking of seasonal cycle is based on the average Pearson’s correlation coefficient for each biome: Good = \(r > 0.80\); Fair = \(r > 0.60\); Poor = \(r < 0.60\); Very poor = \(r < 0\). Ranking of absolute error is based on the average normalized absolute annual error (NAE) for each biome: Low = NAE < 20%; Medium = NAE < 50%; High = NAE < 100%; and Very High = NAE > 100%.

<table>
<thead>
<tr>
<th>Climate</th>
<th>Biome</th>
<th>Seasonal Cycle</th>
<th>Absolute error (annual)</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropics</td>
<td>Evergreen forests</td>
<td>Poor</td>
<td>Low</td>
<td>Very poor (too high)</td>
</tr>
<tr>
<td></td>
<td>Deciduous forests</td>
<td>Poor</td>
<td>Low</td>
<td>Poor (too high)</td>
</tr>
<tr>
<td></td>
<td>Grasslands</td>
<td>Poor</td>
<td>Medium</td>
<td>Very poor (too high)</td>
</tr>
<tr>
<td></td>
<td>Woody savannahs</td>
<td>Fair</td>
<td>Low</td>
<td>Good</td>
</tr>
<tr>
<td>Temperate</td>
<td>Mixed forests</td>
<td>Good</td>
<td>High</td>
<td>Fair</td>
</tr>
<tr>
<td></td>
<td>Grasslands</td>
<td>Good</td>
<td>Low</td>
<td>Poor (too high)</td>
</tr>
<tr>
<td></td>
<td>Woody savannahs</td>
<td>Good</td>
<td>Medium</td>
<td>Fair</td>
</tr>
<tr>
<td></td>
<td>Cropland</td>
<td>Good</td>
<td>Low</td>
<td>Fair</td>
</tr>
<tr>
<td>High latitude</td>
<td>Boreal forests</td>
<td>Good</td>
<td>Medium</td>
<td>Fair</td>
</tr>
<tr>
<td>or altitude</td>
<td>Grasslands</td>
<td>Good</td>
<td>Medium</td>
<td>Fair</td>
</tr>
</tbody>
</table>
Figure 1: Comparison of JULES soil moisture stress factor ($\beta$) to measurements from various potted experiments from Verhoef and Egea (2014). $\beta$ is calculated from Eq. 4. Two different values of $p_0$ (Eq. 5) are shown: $p_0=0.4$ was used for the ‘soil14_p0’ and ‘p0’ soil moisture stress experiments.
Figure 2: Root distributions for JULES with the default 4 layer soil (maximum depth of 3 m), with an updated 14 layer soil (maximum depth of 10.8 m), and compared to parameters from Zeng (2001), where distributions were calculated based on available measurements of root profiles. The parameter $d_e$ in JULES is the e-folding root depth. The plant functional types are: C3, C4 grasses; evergreen and deciduous shrubs (ESH, DSH); needleleaf evergreen trees (NET), temperate broadleaf evergreen trees (BET-Te), broadleaf deciduous trees (BDT), needleleaf deciduous trees (NDT), tropical broadleaf evergreen trees (BET-Tr).
Figure 3. Correlation coefficient for simulated GPP at Fluxnet sites for ten biomes: TrEF = Tropical Evergreen Forests; TrDF = Tropical Deciduous Forests; TrG = Tropical Grasslands; TrS = Tropical Savannahs; TeMF = Temperate Mixed Forests; TeG = Temperate Grasslands; TeS = Temperate Savannahs; Cr = Cropland; CoG = Continental/High altitude grasslands; BoF = Boreal Forests. The sites that fall into each category are listed in the Supplemental Material.
Figure 4. Normalized Absolute Errors for simulated GPP at Fluxnet sites for ten biomes: TrEF = Tropical Evergreen Forests; TrDF = Tropical Deciduous Forests; TrG = Tropical Grasslands; TrS = Tropical Savannahs; TeMF = Temperate Mixed Forests; TeG = Temperate Grasslands; TeS = Temperate Savannahs; Cr = Cropland; CoG = Continental/High altitude grasslands; BoF = Boreal Forests. The sites that fall into each category are listed in the Supplemental Material.
Figure 5: Average seasonal cycle of GPP (gC m$^{-2}$ d$^{-1}$) for representative sites in biomes with large biases. Full dates of simulations are provided in the Supplemental Material, here we give the years included: AU-Fog (2006-2008); BE-Vie (1996-2006); CA-Oas (1996-2010); DE-Tha (1996-2014); IT-Col (1996-2014); LBA-BAN (2004-2006); LBA-K67 (2002-2003); RU-Che (2002-2005); US-Ne1 (2001-2012); US-Ton (2001-2014); ZA-Kru (2000-2013).
Figure 6: Example of impacts of soil moisture stress representations on GPP model skill for two sites: GF-Guy (Tropical evergreen forest), and DE-Tha (Temperate evergreen needleleaf forest). The GF-Guy simulations included years 2007-2009; and the DE-Tha simulations included years 1996-2014. Details of the simulations are provided in Sect. 2 and Tables 1-2.
Figure 7: Example of impacts of various soil moisture stress-related changes (see Table 1) on simulated seasonal cycle of GPP at two sites. The GF-Guy simulations included years 2007-2009; and the DE-Tha simulations included years 1996-2014. Details of the simulations are provided in Sect. 2 and Tables 1-2.