Do land surface models need to include differential plant species responses to drought? Examining model predictions across a latitudinal gradient in Europe

M. G. De Kauwe¹, S.-X. Zhou¹,², B. E. Medlyn¹,³, A. J. Pitman⁴, Y.-P. Wang⁵, R. A. Duursma³, and I. C. Prentice¹,⁶

¹Macquarie University, Department of Biological Sciences, New South Wales 2109, Australia
²CSIRO Agriculture Flagship, Waite Campus, PMB 2, Glen Osmond, SA 5064, Australia
³Hawkesbury Institute for the Environment, University of Western Sydney, Locked Bag 1797, Penrith, NSW, Australia
⁴Australian Research Council Centre of Excellence for Climate Systems Science and Climate Change Research Centre, UNSW, Sydney, Austraila
⁵CSIRO Ocean and Atmosphere Flagship, Private Bag #1, Aspendale, Victoria 3195, Australia
⁶AXA Chair of Biosphere and Climate Impacts, Grand Challenges in Ecosystems and the Environment and Grantham Institute – Climate Change and the Environment, Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot SL5 7PY, UK
Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.
Abstract

Future climate change has the potential to increase drought in many regions of the globe, making it essential that land surface models (LSMs) used in coupled climate models, realistically capture the drought responses of vegetation. Recent data syntheses show that drought sensitivity varies considerably among plants from different climate zones, but state-of-the-art LSMs currently assume the same drought sensitivity for all vegetation. We tested whether variable drought sensitivities are needed to explain the observed large-scale patterns of drought impact. We implemented data-driven drought sensitivities in the Community Atmosphere Biosphere Land Exchange (CABLE) LSM and evaluated alternative sensitivities across a latitudinal gradient in Europe during the 2003 heatwave. The model predicted an overly abrupt onset of drought unless average soil water potential was calculated with dynamic weighting across soil layers. We found that high drought sensitivity at the northernmost sites, and low drought sensitivity at the southernmost sites, was necessary to accurately model responses during drought. Our results indicate that LSMs will over-estimate drought impacts in drier climates unless different sensitivity of vegetation to drought is taken into account.

1 Introduction

Changes in regional precipitation patterns with climate change are highly uncertain (Sillmann et al., 2014), but are widely expected to result in a change in the frequency, duration and severity of drought events (Allen et al., 2010). Drought is broadly defined, but for plants is a marked deficit of moisture in the root zone which results from a period of low rainfall and/or increased atmospheric demand for evapotranspiration. Recently, a series of high-profile drought events (Ciais et al., 2005; Fenshaw et al., 2009; Phillips et al., 2009; Lewis et al., 2011) and associated tree mortality (Breshears et al., 2005; van Mantgem et al., 2009; Peng et al., 2011; Anderegg et al., 2013), have occurred across the globe and these events have led to debate as to whether in-
Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.

Incidence of drought are increasing (Allen et al., 2010; Dai et al., 2013, but see Sheffield et al., 2012). Drought and any ensuing vegetation mortality events have the potential to change land ecosystems from a sink to source (Lewis et al., 2011), and the dominant mechanisms governing the ecosystem responses to drought can vary from reducing stomatal conductance (Xu and Baldocchi, 2003) to increasing tree mortality (Lewis et al., 2010) and changing community species composition (Nepstad et al., 2007).

Our ability to model drought effect on vegetation function is currently limited (Galbraith et al., 2010; Egea et al., 2011; Powell et al., 2013). Remarkably, given the importance of correctly capturing drought impacts on carbon and water fluxes, land surface models (LSMs) designed for use in climate models have rarely been benchmarked against extreme drought events. Mahfouf et al. (1996) compared summertime crop transpiration from 14 land surface schemes, finding that only half of the models fell within the uncertainty range of the observations. They attributed differences among models to the various schemes used by models to represent transpiration processes (e.g. soil water stress function, different number of soil layers) and variability in the initial soil water content at the start of the growing season which relates to variability in the way bare soil evaporation and drainage are represented among different models. Galbraith et al. (2010) showed that a set of dynamic global vegetation models (DGVMs) were unable to capture the 20–30 % reduction in biomass due to drought during a set of throughfall exclusion experiments in the Amazon. Similarly, Powell et al. (2013) demonstrated that a group of five models were unable to predict drought-induced reductions in aboveground biomass (∼ 20 %) in two large-scale Amazon experiments. Galbraith et al. (2010) attributed model variability during drought to: changes in autotrophic respiration, which was not supported by the data, model insensitivity to observed leaf area reductions and the use of different empirical functions to down-regulate productivity during water stress. The models differ both in terms of time-scale of the application of this function (sub-diurnal vs. daily) and whether it is used to down-regulate (net photosynthesis or the maximum rate of Rubisco activity, $V_{\text{max}}$). Gerten et al. (2008) compared the effect of adjusting precipitation regimes on simulated net primary productivity.
(NPP) by four ecosystem models across a range of hydroclimates. They found a consistent direction of change (in terms of NPP) with different scenarios across models but found that the seasonal evolution of soil moisture was differed between the models.

In order for models to better capture realistic responses during drought, they need to draw more closely on experimental data (see Chaves et al., 1993 for a review). One key observation is that there is a continuum of species responses to soil moisture deficit, ranging from isohydric (stomata close rapidly during drought, maintaining a minimum leaf water potential, \( \Psi \)) to anisohydric (stomata remain open during drought, which allows \( \Psi \) to decrease) hydraulic strategies (Tardieu and Simonneau, 1998; Klein, 2014). These differences are widely observed and are thought to be important in determining resilience to drought (McDowell et al., 2008; Mitchell et al., 2013; Garcia-Forner et al., 2015). Many traits, including hydraulic conductivity, resistance to cavitation, turgor loss point, stomatal regulation and rooting depth, contribute to these differences. Systematic differences in the response of leaf gas exchange to soil moisture potential have been observed among species originating from different hydroclimates (Zhou et al., 2013), with species from mesic environments showing stronger stomatal sensitivity to drought than species from xeric environments. Currently, these environmental gradients in species behaviour are not captured in LSMs, which typically assume static plant functional type (PFT) parameterisations. This is in part because historically the data required to describe these attributes have not been available at the global scale, but also due to the necessity of simplification required to run global climate model simulations. PFTs are assumed to have similar or identical sensitivities to drought, with differences only driven by photosynthetic pathway and soil properties. Such an approach ignores experimental evidence of the range of sensitivities to drought among vegetation types (Choat et al., 2012; Limousin et al., 2013; Zhou et al., 2014; Mitchell et al., 2014; Mencuccini et al., 2015), However, it is not known whether observed differences in the response to soil moisture deficit among species are important in determining fluxes at large scales.
In this study we test whether differences in species’ responses to drought are needed to capture drought responses on a continental scale. We built on recent changes to the stomatal conductance ($g_s$) scheme (De Kauwe et al., 2015) within the Community Atmosphere Biosphere Land Exchange (CABLE) LSM (Wang et al., 2011), by implementing a new formulation for drought impacts based on plant ecophysiological studies for 31 species (Zhou et al., 2013, 2014). We obtained three parameterisations for drought response from these studies, characterising low, medium and high sensitivities to drought. We then applied CABLE to simulate responses to an extreme meteorological event, the European 2003 heatwave, at five eddy covariance sites covering a latitudinal gradient. Observations show that there was a significant impact of drought on ecosystem fluxes at these sites (Ciais et al., 2005; Schär et al., 2005). We note that models have been applied to simulate drought effects on productivity (net primary production) and leaf area at individual sites (Ciais et al., 2005; Fischer et al., 2007; Granier et al., 2007; Reichstein et al., 2007) but have not been used to examine whether alternative parameterisations are needed to capture drought responses across sites. We therefore tested how well CABLE was able to simulate the impact of drought on carbon and water fluxes at these sites using alternative parameterisations for drought sensitivity. We hypothesised that drought sensitivity would increase with latitude, and accounting for that latitudinal gradient in drought sensitivity would improve the performance of CABLE.

2 Methods

2.1 Model description

CABLE represents the vegetation using a single layer, two-leaf canopy model separated into sunlit and shaded leaves (Wang and Leuning, 1998), with a detailed treatment of within canopy turbulence (Raupach, 1994; Raupach et al., 1997). Soil water and heat conduction is numerically integrated over six discrete soil layers following the
Richards equation and up to three layers of snow can accumulate on the soil surface. A complete description can be found in Kowalczyk et al. (2006) and Wang et al. (2011). CABLE has been used extensively for both offline (Abramowitz et al., 2008; Wang et al., 2011; De Kauwe et al., 2015) and coupled simulations (Cruz et al., 2010; Pitman et al., 2011; Mao et al., 2011; Lorenz et al., 2014) within the Australian Community Climate Earth System Simulator (ACCESS, see http://www.accessimulator.org.au; Kowalczyk et al., 2013); a fully coupled earth system model. The source code can be accessed after registration at https://trac.nci.org.au/trac/cable.

2.2 Representing drought stress within CABLE

We build on the work by De Kauwe et al. (2015), who introduced a new \( g_s \) scheme into CABLE. In this scheme, stomata are assumed to behave optimally (Cowan and Farquhar, 1977) leading to the following formulation of \( g_s \) (Medlyn et al., 2011)

\[
g_s = g_0 + 1.6 \left( 1 + \frac{g_1}{\sqrt{D}} \right) \frac{A}{C_s} \tag{1}
\]

where \( A \) is the net assimilation rate (\( \mu \text{mol m}^{-2} \text{s}^{-1} \)), \( C_s \) (\( \mu \text{mol mol}^{-1} \)) and \( D \) (kPa) are the CO\(_2\) concentration and the vapour pressure deficit at the leaf surface, respectively, and \( g_0 \) (\( \text{mol m}^{-2} \text{s}^{-1} \)), and \( g_1 \) are fitted constants representing the residual stomatal conductance as \( A \) rate reaches zero, and the slope of the sensitivity of \( g_s \) to \( A \), respectively. The model was parameterised for different PFTs using data from Lin et al. (2015) (see De Kauwe et al., 2015).

In the standard version of CABLE, drought stress is implemented as an empirical scalar (\( \beta \)) that depends on soil moisture content, weighted by the fraction of roots in each of CABLE’s six soil layers:

\[
\beta = \sum_{i=1}^{n} f_{\text{root},i} \frac{\theta_i - \theta_w}{\theta_{fc} - \theta_w}; \quad \beta \in [0,1] \tag{2}
\]
where $\theta_i$ is the volumetric soil moisture content ($m^3 m^{-3}$) in soil layer $i$, $\theta_w$ is the wilting point ($m^3 m^{-3}$), $\theta_{fc}$ is the field capacity ($m^3 m^{-3}$) and $f_{\text{root},i}$ is the fraction of root mass in soil layer $i$. The six soil layers in CABLE have depths 0.022, 0.058, 0.154, 0.409, 1.085 and 2.872 m. The factor $\beta$ is assumed to limit the slope of the relationship between stomatal conductance ($g_s$, mol m$^{-2}$ s$^{-1}$; Leuning, 1995) by acting as a modifier on the parameter $g_1$.

In this study, we introduced a new expression for drought sensitivity of gas exchange, based on the work of Zhou et al. (2013, 2014). In this model, both $g_1$ and the photosynthetic parameters $V_{cmax}$ and $J_{max}$ are assumed to be sensitive to pre-dawn leaf water potential, but this sensitivity varies across species. There is considerable evidence that both $g_1$ and $V_{cmax}$ are sensitive to soil moisture (Keenan et al., 2009; Egea et al., 2011; Flexas et al., 2012; Zhou et al., 2013). There is also widespread evidence that plants are more directly respond to water potential rather than water content (Comstock and Mencuccini, 1998; Verhoef and Egea, 2014).

Zhou et al. (2013) extended the optimal stomatal model of Medlyn et al. (2011) by fitting an exponential function to relate $g_1$ to pre-dawn leaf water potential ($\Psi_{pd}$):

$$g_1 = g_{1\text{wet}} \times \exp(b \Psi_{pd})$$

where $g_{1\text{wet}}$ is fitted parameter representing plant water use under well watered conditions (i.e. when $\Psi_{pd} = 0$) and $b$ is a fitted parameter representing the sensitivity of $g_1$ to $\Psi_{pd}$. Species with different water use strategies can by hypothesised to differ in not only their $g_1$ parameter under well-watered conditions, $g_{1\text{wet}}$ (see Lin et al., 2015), but also with the sensitivity to $\Psi_{pd}$, $b$. Zhou et al. (2013) also advanced a non-stomatal limitation to the photosynthetic biochemistry, which describes the apparent effect of water stress on $V_{cmax}$:

$$V_{cmax} = V_{cmax, \text{wet}} \frac{1 + \exp(S_f \Psi_f)}{1 + \exp(S_f (\Psi_f - \Psi_{pd}))}$$
where $V_{c_{\text{max}},\text{wet}}$ is the $V_{c_{\text{max}}}$ value in well watered conditions, $S_f$ is a sensitivity parameter describing the steepness of the decline with water stress, $\Psi_f$ is the water potential at which $\Psi_{pd}$ decreases to half of its maximum value. As with $g_1$, it is hypothesised that in the same way species vary in their $V_{c_{\text{max}}}$ values in well-watered conditions ($V_{c_{\text{max}},\text{wet}}$), they would also differ in their sensitivity of down-regulated $V_{c_{\text{max}}}$ with water stress (Zhou et al., 2014). In CABLE, as there is a constant ratio between the parameters $J_{\text{max}}$ and $V_{c_{\text{max}}}$, the parameter $J_{\text{max}}$ is similarly reduced by drought.

To implement Eq. (6) in CABLE we first had to convert soil moisture content ($\theta$) to pre-dawn leaf water potential ($\Psi_{pd}$). We did so by assuming that overnight $\Psi_{pd}$ and $\Psi_S$ equilibrate before sunrise, thus ignoring any night-time transpiration (Dawson et al., 2007). Following Cambell (1974), we related $\theta$ to $\Psi_S$ in each soil layer by:

$$\Psi_{S,i} = \Psi_e \left( \frac{\theta_i}{\theta_{\text{sat}}} \right)^{-k}$$

(5)

where $\Psi_e$ is the air entry water potential (MPa) and $k$ (unitless) is an empirical coefficient which is related to the soil texture. Values for $\Psi_e$ and $b$ are taken from CABLE’s standard lookup table following Clapp and Hornberger (1978). We then needed to obtain a representative weighted estimate of $\Psi_S$ across CABLE’s soil layers. We tested three potential approaches for weighting in this paper:

1. Using the root-biomass weighted $\theta$ and converting this to $\Psi_S$ using Eq. (8).
2. Taking the integrated $\theta$ over the top 5 soil layers (1.7 m depth) and converting this to $\Psi_S$ using Eq. (8). This method assumes the plant effectively has access to an entire “bucket” of soil water.
3. Weighting the average $\Psi_S$ for each of the six soil layers by the weighted soil-to-root conductance to water uptake of each layer, following Williams et al. (1996, 2001). The total conductance term depends the combination of a soil component
(R_S) and a root component (R_r). R_S is defined as Gardner (1960):

\[ R_S = \frac{\ln\left(\frac{r_s}{r_f}\right)}{2\pi r_f D G_{\text{soil}}} \] (6)

where \( r_s \) is the mean distance between roots (m), \( r_f \) is the fine root radius (m), \( D \) is the depth of the soil layer, \( G_{\text{soil}} \) is the soil conductivity (mmol m\(^{-1}\) s\(^{-1}\) MPa\(^{-1}\)) which depends on soil texture and soil water content, \( l_r \) is the fine root density (mm\(^{-3}\)). \( R_r \) is defined as:

\[ R_r = \frac{R_r^*}{F D} \] (7)

where \( R_r^* \) is the root resistivity (MPa s g mmol\(^{-1}\)), \( F \) is the root biomass per unit volume (g m\(^{-3}\)). This method weights \( \Psi_S \) to the upper soil layers when the soil is wet, but shifts towards layer lowers as the soil dries, due to the lower soil hydraulic conductance.

### 2.3 Model simulations

During 2003, Europe experienced an anomalously dry summer, amplified by a combination of a preceding dry spring and high summer temperatures (Ciais et al., 2005; Schär et al., 2005). Summer temperatures were recorded to have exceeded the 30 year June–July–August (JJA) average by 3 °C (Schär et al., 2005). Consequently we choose to focus our model comparisons on this period, in particular the period between June and September 2003.

At each of the five Fluxnet sites we ran three sets of simulations:

- a control simulation ("CTRL"), representing CABLE version 2.0.1.
- Three simulations to explore the new drought model using a “high” (Quercus robur), “medium” (Quercus ilex) and “low” (Cedrus atlantica) sensitivity to...
soil moisture. Parameter values were obtained from the meta-analysis by Zhou et al. (2013, 2014) and are given in Table 1. For each of these simulations we also tested the three different methods of obtaining $\Psi_S$ as described above.

- A “no drought” simulation in which any transpired water was returned to the soil. By comparing this simulation with either the control or any of the new drought model simulations (high, medium, low), a guide to the magnitude of the drought should be apparent.

Model parameters were not calibrated to match site characteristics; instead default PFT parameters were used for each site. Although CABLE has the ability to simulate full carbon, nitrogen and phosphorus biogeochemical cycling, this feature was not activated for this study, instead only the carbon and water cycle were simulated. For all simulations, leaf area index (LAI) was prescribed using CABLE’s gridded monthly LAI climatology derived from Moderate-resolution Imaging Spectroradiometer (MODIS) LAI data (Knyazikhin et al., 1998, 1999) and the $g_s$ scheme following Medlyn et al. (2011; see De Kauwe et al., 2015) was used throughout. All model simulations were spun-up by repeating the meteorological forcing site data until soil moisture and soil temperatures reached equilibrium (as we were ignoring the full biogeochemical cycling in these simulations).

### 2.3.1 Water use efficiency bug

During the course of these simulations we identified a bug in the way CABLE calculates carbon and water fluxes during drought. If CABLE is unable to meet atmospheric demand for water, it down-regulates transpiration to match the available supply from the soil; however, CABLE does not in turn down-regulate photosynthesis sufficiently to match this reduced water flux. Consequently, during periods of extreme drought CABLE’s decoupling of the water and carbon cycles, results in very high water use efficiency (WUE). Fixing this issue is not straightforward because of the way the model solves for multiple components at once (e.g. $C_i$, $A_n$, $D$, leaf temperature, etc.) and
is thus beyond the scope of this work. Crucially, the issue only relates to an over-
estimation of photosynthesis during drought; the transpiration and latent heat fluxes are
both calculated correctly and thus does not preclude useful assessment of the standard
CABLE model. The issue arises because the estimated $\beta$ using Eq. (2) does not suffi-
ciently limit carbon and water fluxes during periods of water stress. In the new drought
model the drought sensitivities are stronger, and so we do not reach this scenario for
which root water supply is insufficient to meet demand (requiring down-regulation of
transpiration).

2.4 Datasets used

To assess the performance of the CABLE model both with and without the new drought
scheme, we selected a gradient of five forested Fluxnet (http://www.fluxdata.org/) sites
across Europe (Table 2). Site data were obtained via the Protocol for the Analysis of
Land Surface models (PALS; Abramowitz, 2012). These data have previously been pre-
processed and quality controlled for use within the LSM community. Consequently, all
site-years had near complete observations of key meteorological drivers (as opposed
to significant gap-filled periods).

3 Results

Table 3 summarises summer differences in rainfall, air temperature, GPP and LE be-
tween 2002 and 2003 across the five sites covering a latitudinal gradient across Eu-
rope. Whilst the impact of the 2003 heatwave varied between sites, every site was
warmer and drier in 2003. Similarly, GPP was lower at every site except Espirra and LE
was lower at three of the sites (Hesse, Roccarespampani and Castelporziano). Figure 1
shows a site-scale comparison between standard CABLE (CTRL) and the observed
GPP, LE and transpiration ($E$) at the five sites. Tables 4 and 5 show a series of sum-
mary statistics (Root Mean Squared Error (RMSE), Nash–Sutcliffe efficiency (NSE),
Pearson’s correlation coefficient \( (r) \) between modelled and observed LE and GPP. An indication of the severity of the drought can be obtained by looking at the difference between the “No drought” and the CTRL simulation. For the two northernmost sites (Tharandt and Hesse), the CTRL simulation, the CTRL simulation generally matches the trajectory of the observed LE, but displays systematic periods of over-estimation (i.e. under-estimating the drought effect). By contrast, in the three southernmost sites (Roccabruna, Castelporziano and Espirra), the reverse is true: the CTRL simulations descend into drought stress much more quickly than the observed fluxes. This rapid drought progression is particularly evident around day of year 155 at the Roccabruna site. Across all sites, the CTRL simulations agree more closely with the observed fluxes in terms of GPP (NSE = −0.64 to −0.61; \( r = 0.71 \) to 0.84), rather than LE (NSE = −8.55 to −0.04; \( r = 0.14 \) to 0.70). The better prediction of GPP than LE is partly explained by the WUE bug that we identified in the methods section (Sect. 2.3). The CTRL simulations over-predict the sensitivity of the LE fluxes to drought, but do not down-regulate the GPP at the same time, resulting in closer agreement with the observed for GPP (essentially compensating for CABLE’s over-sensitivity of GPP to drought). The second part of the explanation relates to soil evaporation. Around mid-spring, CABLE simulates very large LE fluxes that are not evident in the \( E \) fluxes. This over-evaporation of soil water results in model-observation divergence in LE fluxes, which is not evident in the model-observation GPP comparison.

We now consider the implementation of the new drought model and the three sensitivity parameterisations. Figure 2a shows how leaf-level photosynthesis is predicted to decline (using Eqs. 3 and 4) in the new drought model with increasing water stress (more negative \( \Psi_s \)). The different sensitivities to drought are clearly visible, with the three parameterisations representing a spectrum of behaviour ranging from high to low drought sensitivity. Figure 2b and c shows how the new drought model compares to the standard CABLE (CTRL; using Eq. 2) model on a sandy and clay soil type. The CTRL model is seen to most closely match the high sensitivity simulation on a sandy soil, but it predicts an earlier descent into drought stress. By contrast on the clay soil, the new
medium and high sensitivity simulations encompass the predictions from the CTRL model. The new drought model and parameterisations afford a more flexible sensitivity to the down-regulation of photosynthesis with drought, which is particularly evident in the low sensitivity simulation.

Figures 3–7 show the same site comparisons as Fig. 1, but with the addition of the new drought model and the three different ways in which \( \Psi_S \) can be averaged over the soil profile. Across all sites it is clear that when using the first method of deriving \( \Psi_S \) the new drought model behaves in much the same way as the CTRL simulation. The explanation is that weighting \( \Psi_S \) by the fraction of roots in a layer results in water being principally extracted from the top three shallow layers (Supplement Figs. S1–S5). Consequently, small changes in \( \theta \) result in a rapid decline in \( \Psi_S \) (owing to the non-linear relationship between \( \theta \) and \( \Psi_S \), Fig. 1), which results in an unrealistically abrupt shutdown of transpiration. Method two for deriving \( \Psi_S \) shows a greater separation between the three sensitivity parameterisations than method one. The greater separation is most evident at the southern sites; the model performs particularly well at Espirra (LE RMSE < 16 vs. CTRL RMSE = 41.72) and to a lesser extent at Castelporziano (LE low sensitivity RMSE = 19.18 vs. CTRL RMSE = 31.11). Nevertheless, at the two northern sites, the model completely underestimates the size of the drought, as a result of using a large soil water bucket (1.7 m) to calculate \( \Psi_S \). The third method of deriving \( \Psi_S \) in combination with the new drought model generally works the best across all the sites, as it allows CABLE to simulate a more gradual reduction of fluxes during drought. At Roccarespampani a medium drought sensitivity performs best at reproducing the observed LE (CTRL RMSE = 38.15 vs. 18.41), whilst at Castelporziano (CTRL RMSE = 31.11 vs. 20.37) and Espirra (CTRL RMSE = 41.72 vs. 15.41) the low sensitivity (or a medium sensitivity at Castelporziano (LE RMSE = 20.46) works best. By contrast, at the two northernmost sites, a high drought sensitivity works best (Hesse LE CTRL RMSE = 31.57 vs. 28.72); although at Tharandt the new drought model performs marginally worse than the CTRL (LE CTRL RMSE = 22.0 vs. 25.48).
4 Discussion

Experimental data suggest that plants exhibit a continuum of drought sensitivities, with species originating in more mesic environments showing higher sensitivity than species from more xeric environments. We investigated whether variable drought sensitivity improves the ability of the CABLE LSM to reproduce observed drought impacts across a latitudinal gradient. We found that, at the northernmost sites, a high drought sensitivity was required; moving southwards, the sensitivity parameterisation transitioned to a medium and finally to a low drought sensitivity. This work demonstrates the importance of understanding how plant traits vary with climate across the landscape. However, our analysis also highlighted the importance of identifying which soil layers matter most to the plant: our results depended strongly on how we weighted soil moisture availability through the profile.

4.1 Weighting soil moisture availability

Commonly, empirical dependences of gas exchange on soil moisture content or potential (Eqs. 3 and 4) are estimated from pot experiments (e.g. Zhou et al., 2013, 2014), in which it is fair to assume that the soil moisture content is uniform. In contrast, soil moisture content in the field typically has a strong vertical profile. Thus, to implement such equations in a land surface model requires that we specify how to weight the soil layers to obtain a representative value of whole-profile \( \theta \) or \( \Psi_S \). In this study we tested three potential implementations. Our first approach was to weight each layer by root biomass. Evidence suggests that plants preferentially access regions in the root zone where water is most freely available (Green and Clotheir, 1995; Huang et al., 1997). Hence, many models follow this approach: for example, the original version of CABLE weighted soil moisture content by root biomass (Eq. 2) while the Community Land Model (CLM) estimates a water stress factor based on a root-weighted \( \Psi_S \), using a PFT-defined minimum and maximum water potential (Oleson et al., 2013). However, we found that this approach performed poorly. We observed an “on–off” behaviour in
response to drought, which occurs because the behaviour of the model is driven by the top soil layers, whose total soil moisture content is relatively small and root biomass is relatively high, and can be depleted rapidly, leading to a sudden onset of severe drought. Many other LSMs show this abrupt effect of drought (Egea et al., 2011; Powell et al., 2013). Powell et al. (2013) found that four models (CLM version 3.5, Integrated Biosphere Simulator version 2.6.4 (IBIS), Joint UK Land Environment Simulator version 2.1 (JULES), and Simple Biosphere model version 3 (SiB3)) implement abrupt transitions of this kind. We also found that with this weighting of soil layers, there was little effect of variable drought sensitivity: the depletion of soil moisture content of the top layers is so rapid that there is little difference between low and high sensitivities to drought. Such an outcome suggests that there is little adaptive significance of drought sensitivity, which seems unlikely. A further implication of using a root-weighted function to calculate $\Psi_S$ is that two distinctly different scenarios, a soil that has been very wet but experienced a short dry period, allowing the topsoil to dry, and a soil that has had a prolonged period of drought but experienced a recent rainfall event, would have similar impacts on gas exchange. Again, this outcome seems unlikely.

We tested a second implementation in which soil moisture potential was calculated from the moisture content of the entire rooting zone (top five soil layers = 1.7 m). Such an approach is commonly used in forest productivity models (e.g. Landsberg and Waring, 1997). However, this approach severely underestimates drought impacts because the moisture content of the total soil profile is so large, meaning that it is rarely depleted enough to impact on gas exchange.

In reality, plant water uptake shifts lower in the profile as soil dries out (e.g. Duursma et al., 2011). Thus, in our third implementation, we tested an approach in which the weighting of soil layers moves downwards as drought progresses. This approach is effectively similar to that used by the soil–plant–atmosphere (SPA) model (Williams et al., 1996, 2001), in which soil layers are weighted by their soil-to-root conductance, which declines as the moisture content declines. This approach performed best of the three we tried, and allowed CABLE to replicate the observations across the latitudi-
nal gradient. This dynamic weighting of $\Psi_S$ may partially explain previous good performance by SPA in other model inter-comparisons focussed on drought (e.g. Powell et al., 2013). Recently, Bonan et al. (2014) tested the suitability of using a model that considers optimal stomatal behaviour and plant hydraulics (SPA; Williams et al., 1996) for earth system modelling, and demonstrated marked improvement over the standard model during periods of drought stress. We thus suggest that models using a soil moisture stress function to simulate drought effects on gas exchange should consider a dynamic approach to weighting the contribution of different soil layers.

We note that this issue is related to another long-standing problem for LSMs: that of determining the vertical distribution of root water uptake (e.g. Feddes et al., 2001; Federer et al., 2003; Kleidon and Heimann, 1998, 2000). In the standard version of CABLE, water uptake from each soil layer initially depends on the fraction of root biomass in each layer, but moves downwards during drought as the upper layers are depleted. It is possible that changes to the weighting of soil moisture in determining drought sensitivity should also be accompanied by changes to the distribution of root water uptake, but we did not explore this option here. Li et al. (2012) previously tested an alternative dynamic root water uptake function (Lai and Katul, 2000) in CABLE, but found little improvement in predicted LE during seasonal droughts without also considering a mechanism for hydraulic redistribution. Further work should evaluate models not only against LE fluxes, but also against measurements of soil moisture profiles. Many experimental sites now routinely install multiple soil moisture sensors (e.g. direct gravimetric sampling, neutron probes, time domain reflectometry), which provide accurate insight into root water extraction and hydraulic redistribution, even down to considerable depths (> 4 m). These data have thus far been underutilised for model improvement, but should be a priority for reducing the uncertainty in soil moisture dynamics.
4.2 Incorporating different sensitivities to drought

Using the third and best method to calculate overall \( \Psi_s \), we found that varying drought sensitivity across sites enabled the model to better capture drought effects across the latitudinal gradient, with a high drought sensitivity implied in northern sites and a low drought sensitivity implied in southern sites. These results should not be surprising, given the increasing amount of experimental evidence suggesting that drought sensitivity varies among species and across climates (e.g. Engelbrecht and Kursar, 2003; Engelbrecht et al., 2007; Skelton et al., 2015). In contrast to these data, most LSMs assume a single parameterisation for drought sensitivity, which is typically based on mesic vegetation. Our results suggest that such a parameterisation is very likely to overstate the impacts of drought in drier regions.

Our work thus underlines a need to move beyond models that implement drought sensitivity through a single PFT parameterisation. In order to capture the observed variability in plant responses to drought, models need to consider a continuum of sensitivities. It is, of course, challenging to implement such a continuum in a global vegetation model. In this study, we used a simple site-specific approach in which we selected three sets of model parameters from a meta-analysis by Zhou et al. (2013, 2014), allowing us to characterise a range of plant responses to drought. Global vegetation models would require a more sophisticated approach that relates drought sensitivity to the climate of each pixel. One potential solution would be to develop an empirical correlation between drought sensitivity and a long-term moisture index (e.g. the ratio of mean precipitation to the equilibrium evapotranspiration; Cramer and Prentice, 1988; Gallego-Sala et al., 2010). Previous studies have demonstrated the feasibility of linking model parameters that determine plant water use strategy to such a moisture index in global simulations (Wang et al., 2014; De Kauwe et al., 2015). Such an approach would require a concerted effort to collate appropriate data, as there are few compilations to date of traits related to drought sensitivity (but see Manzoni et al., 2011; Zhou et al., 2013). Another, more challenging, alternative, would be to develop optimization
4.3 Further model uncertainties

Whilst this work advances the ability of LSMs to simulate drought, it does not address all processes needed to correctly capture drought impacts. Other issues to consider include: (i) rooting depth, (ii) leaf shedding, (iii) soil evaporation, and (iv) soil heterogeneity, among others.

Here we have assumed that all sites had the soil depth (4.6 m), with rooting depth distributed exponential through the profile, as is commonly used in LSMs. However, this assumption may be incorrect. Access to water by deep roots could be a potential alternative explanation for the low drought sensitivity that we inferred at the southernmost site, Espirra. Here the dominant species is not native to the region, but rather a plantation of blue gum (Eucalyptus globulus), a species that is generally found to have high, not low, drought sensitivity (White, 1996; Mitchell et al., 2014). Many eucalypts have a deep rooting strategy (Fabiao et al., 1987), suggesting a possible alternative explanation for drought tolerance at this site. More in-depth study of fluxes and soil moisture patterns at this site would be needed to determine the role of rooting depth.

During droughts, plants are often observed to shed their leaves as a self-regulatory mechanism to reduce water losses (Tyree et al., 1993; Jonasson et al., 1997; Bréda et al., 2006). During the 2003 heatwave, at Hesse an early reduction of approximately 1.7 \( m^2 m^{-2} \) was observed, similarly at Brasschaat there was a observed reduction of 0.8 \( m^2 m^{-2} \) and at Tharandt needle-litter was increased during September until November, with LAI estimated to be 0.9 \( m^2 m^{-2} \) lower (Bréda et al., 2006; Granier et al., 2007). In contrast, models typically fix turnover rates for leaves and as such this feedback is largely absent from models. During periods of water stress, models do simulate an indirect reduction in LAI via down-regulated net primary productivity; however this feedback is much slower than is commonly observed. Not accounting for the canopy scale feedback will result in models over-estimating carbon and water fluxes and thus losses in \( \theta \).
during drought. Thus, models will also over-predict their sensitivity to drought because of their reliance on $\theta$ to determine drought status.

Existing models also disagree as to the mechanism by which to down-regulate productivity during periods of water stress (De Kauwe et al., 2013). In the standard version of CABLE, only the slope of the relationship between $g_s$ and $A$ is reduced by water stress. The SPA model behaves similarly. In contrast, JULES (Clark et al., 2011) and the Sheffield Dynamic Global Vegetation Model (SDGVM; Woodward and Lomas, 2004), down-regulate the photosynthetic capacity via the biochemical parameters $V_{c_{\text{max}}}$ and $J_{\text{max}}$ (maximum electron transport rate). Here, we assumed that water stress affects both the slope of $g_s - A$ and the biochemical parameters $V_{c_{\text{max}}}$ and $J_{\text{max}}$, supported by results from Zhou et al. (2013, 2014). We did not evaluate this assumption against the eddy flux data. However, previous studies have also suggested that both effects are needed to explain responses of fluxes during drought (Keenan et al., 2010).

Losses via soil evaporation play an important role in determining drought depth, particularly at sites with low LAI. Here we identified an early-season over-estimation of soil evaporation in CABLE that was apparent across all of the European flux tower sites. Alternative soil schemes have been tested within CABLE, for example the SLI soil model adjusts the soil boundary layer resistance and adds a litter layer (Haverd and Cuntz, 2010). A new hydrological model has been developed by Decker, 2015 (submitted) that includes lateral flows, sub-grid scale soil moisture variability, groundwater and a replacement formulation of soil evaporation. This improved the simulation of total evaporation in CABLE but still lacks the drought responses discussed here and therefore is unlikely to solve the problems identified. Overall, it is likely that LSMs will require a simultaneous development of the hydrological and ecological parameterizations, something that is unfortunately rare in land surface modelling where the objectives are linked with global climate modelling.

Finally, although models do have the capacity to simulate vertical variations in $\theta$, they do not always represent any horizontal sub-grid scale variability. This assumption is likely to contribute to the abruptness of modelled transitions from well-watered to
completely down-regulated carbon and water fluxes. Earlier work by Entekhabi and Eagleson (1989), and models such as the variable infiltration capacity (VIC) model (Liang et al., 1994), and most recently Decker, 2015 (submitted) have attempted to address this issue by employing statistical distributions to approximate horizontal spatial heterogeneity in soil moisture (see also Crow and Wood, 2002). These parsimonious approaches typically require few parameters, making them attractive in the LSM context and potentially suitable for modelling ecosystem and hydrological responses to drought (Luo et al., 2013).

### 4.4 Testing models against extreme events

We have used a model evaluation against flux measurements during a large-scale heatwave event to make significant progress in modelling of drought impacts. While model evaluation against data is now commonplace (Prentice et al., 2015) and has recently been extended to formal benchmarking, particularly in the land surface community (Abramowitz, 2005; Best et al., 2015), many of these benchmarking indicators are based on seasonal or annual outputs and thus miss the opportunity to examine model performance during extreme events. Furthermore, the key weakness of model evaluation in its current inception is the focus on tests that aim to identify the best overall model performance; this applies particularly to land surface schemes. This is something the benchmarking approach used by Abramowitz (2005) and Best et al. (2015) seek to avoid by testing model performance against empirical benchmarks. Overall however, there remains a tendency to trade mechanistic realism is often traded for present day accuracy, with limited real improvement in model structure. Model projections under future climate change should require that we have good mechanistic representations of the impacts of extreme events. However, this is hardly ever tested and there is therefore an urgent need to orient model testing to periods of extremes. To that end, precipitation manipulation experiments (e.g. Nepstad et al., 2002; Hanson et al., 2003; Pangle et al., 2012) represent a good example of a currently under-exploited avenue (but see Fisher et al., 2007; Powell et al., 2013) that could be used
for model evaluation and/or benchmarking (Smith et al., 2014). However, we urge that these exercises do not focus solely on overall model performance, but also test the realism of individual model assumptions (Medlyn et al., 2015).

The Supplement related to this article is available online at doi:10.5194/bgd-12-12349-2015-supplement.

Acknowledgements. This work was supported by the Australian Research Council (ARC) Linkage grant LP140100232 and the ARC Centre of Excellence for Climate System Science (CE110001028). S.-X. Zhou was supported by an international Macquarie University Research Excellence Scholarship. This work is also a contribution to the AXA Chair Programme in Biosphere and Climate Impacts and the Imperial College initiative on Grand Challenges in Ecosystems and the Environment. We thank CSIRO and the Bureau of Meteorology through the Centre for Australian Weather and Climate Research for their support in the use of the CABLE model. This work used eddy covariance data acquired by the FLUXNET community for the La Thuile FLUXNET release, supported by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Universite Laval and Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California, Berkeley, University of Virginia. All data analysis and plots were generated using the Python language and the Matplotlib plotting library (Hunter, 2007).
References


Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.


Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.


Do land surface models need to include differential plant species responses?

M. G. De Kauwe et al.


Do land surface models need to include differential plant species responses?

M. G. De Kauwe et al.


Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.


Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.


Table 1. Baseline parameter values used to represent the three sensitivities: “high” (*Quercus robur*), “medium” (*Quercus ilex*) and “low” (*Cedrus atlantica*) to drought stress. Parameter values are taken from Zhou et al. (2013, 2014).

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>$b$</th>
<th>$S_f$</th>
<th>$\Psi_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>1.55</td>
<td>6.0</td>
<td>−0.53</td>
</tr>
<tr>
<td>Medium</td>
<td>0.82</td>
<td>1.9</td>
<td>−1.85</td>
</tr>
<tr>
<td>Low</td>
<td>0.46</td>
<td>5.28</td>
<td>−2.31</td>
</tr>
</tbody>
</table>
Table 2. Summary of flux tower sites.

<table>
<thead>
<tr>
<th>Site</th>
<th>PFT</th>
<th>Dominant species</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Country</th>
<th>Sand/Silt/Clay Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tharandt</td>
<td>ENF</td>
<td>Picea abies</td>
<td>50°58’ N</td>
<td>13°34’ E</td>
<td>Germany</td>
<td>0.37/0.33/0.3</td>
</tr>
<tr>
<td>Hesse</td>
<td>DBF</td>
<td>Fagus sylvatica</td>
<td>48°40’ N</td>
<td>7°05’ E</td>
<td>France</td>
<td>0.37/0.33/0.3</td>
</tr>
<tr>
<td>Roccarespampani</td>
<td>DBF</td>
<td>Quercus cerris</td>
<td>42°24’ N</td>
<td>11°55’ E</td>
<td>Italy</td>
<td>0.6/0.2/0.2</td>
</tr>
<tr>
<td>Castelporziano</td>
<td>EBF</td>
<td>Quercus ilex</td>
<td>41°42’ N</td>
<td>12°22’ E</td>
<td>Italy</td>
<td>0.6/0.2/0.2</td>
</tr>
<tr>
<td>Espirra</td>
<td>EBF</td>
<td>Eucalyptus globulus</td>
<td>38°38’ N</td>
<td>8°36’ W</td>
<td>Portugal</td>
<td>0.37/0.33/0.3</td>
</tr>
</tbody>
</table>
Table 3. Mean change in climate and fluxes between 2002 and 2003 covering the period between June and September.

<table>
<thead>
<tr>
<th>Site</th>
<th>Precipitation (mm month$^{-1}$)</th>
<th>Air temperature ($^\circ$C)</th>
<th>GPP (g C m$^{-2}$ month$^{-1}$)</th>
<th>LE (W m$^{-2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tharandt</td>
<td>$-115.57$</td>
<td>1.45</td>
<td>$-38.45$</td>
<td>0.52</td>
</tr>
<tr>
<td>Hesse</td>
<td>$-49.20$</td>
<td>2.98</td>
<td>$-123.38$</td>
<td>$-11.90$</td>
</tr>
<tr>
<td>Roccarespampani</td>
<td>$-87.36$</td>
<td>2.18</td>
<td>$-71.94$</td>
<td>$-6.17$</td>
</tr>
<tr>
<td>Castelporziano</td>
<td>$-20.31$</td>
<td>4.57</td>
<td>$-49.73$</td>
<td>$-6.47$</td>
</tr>
<tr>
<td>Espira</td>
<td>$-14.45$</td>
<td>1.77</td>
<td>28.46</td>
<td>22.83</td>
</tr>
</tbody>
</table>
Table 4. Summary statistics of modelled and observed LE at the five FLUXNET sites during the main drought period (1 June–31 August 2003). For each site the best performing model simulation has been highlighted in bold.

<table>
<thead>
<tr>
<th>Site</th>
<th>$\Psi$ Method</th>
<th>Root Mean Squared Error (RMSE)</th>
<th>Nash–Sutcliffe efficiency (NSE)</th>
<th>Pearsons’s correlation coefficient ($r$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Tharandt</td>
<td>1</td>
<td>22.0</td>
<td>23.77</td>
<td>25.50</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>34.07</td>
<td>35.88</td>
<td>36.77</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25.48</td>
<td>29.05</td>
<td>32.13</td>
</tr>
<tr>
<td>Hesse</td>
<td>1</td>
<td>31.57</td>
<td>34.48</td>
<td>38.93</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>51.29</td>
<td>58.25</td>
<td>62.15</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>28.72</td>
<td>44.96</td>
<td>56.26</td>
</tr>
<tr>
<td>Roccarespampiani</td>
<td>1</td>
<td>38.15</td>
<td>47.23</td>
<td>43.43</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>30.69</td>
<td>24.52</td>
<td>21.85</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>45.09</td>
<td>18.41</td>
<td>29.48</td>
</tr>
<tr>
<td>Castelporziano</td>
<td>1</td>
<td>31.11</td>
<td>38.41</td>
<td>40.85</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>30.53</td>
<td>28.22</td>
<td>19.18</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>39.08</td>
<td>20.46</td>
<td>20.37</td>
</tr>
<tr>
<td>Espira</td>
<td>1</td>
<td>41.72</td>
<td>41.27</td>
<td>41.10</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15.87</td>
<td>13.96</td>
<td>13.94</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>40.79</td>
<td>20.47</td>
<td>15.41</td>
</tr>
</tbody>
</table>
Table 5. Summary statistics of modelled and observed GPP at the five FLUXNET sites during the main drought period (1 June–31 August 2003). For each site the best performing model simulation has been highlighted in bold.

<table>
<thead>
<tr>
<th>Site</th>
<th>$W_s$ Method</th>
<th>Root Mean Squared Error (RMSE)</th>
<th>Nash–Sutcliffe efficiency (NSE)</th>
<th>Pearson’s correlation coefficient ($r$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CTRL High Medium Low</td>
<td>CTRL High Medium Low</td>
<td>CTRL High Medium Low</td>
</tr>
<tr>
<td>Tharandt</td>
<td>1</td>
<td>2.16; 2.04; 2.09</td>
<td>0.38</td>
<td>0.27; 0.35; 0.32</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.25; 2.28; 2.30</td>
<td>0.21; 0.18; 0.17</td>
<td>0.71; 0.52; 0.51</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.20; 2.11; 2.19</td>
<td>0.24; 0.30; 0.25</td>
<td>0.73; 0.66; 0.61</td>
</tr>
<tr>
<td>Hesse</td>
<td>1</td>
<td>3.37; 2.45; 2.90</td>
<td>0.54</td>
<td>0.31; 0.61; 0.46</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.70; 3.23; 3.46</td>
<td>0.53; 0.53; 0.23</td>
<td>0.76; 0.80; 0.79</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.42; 2.69; 3.24</td>
<td>0.24; 0.46; 0.33</td>
<td>0.84; 0.75; 0.66</td>
</tr>
<tr>
<td>Roccarespampini</td>
<td>1</td>
<td>3.45; 2.58; 2.12</td>
<td>0.58</td>
<td>−0.12; 0.38; 0.58</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.83; 1.30; 2.70</td>
<td>0.69; 0.84; 0.31</td>
<td>0.84; 0.94; 0.88</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.72; 1.74; 3.08</td>
<td>−0.30; 0.72; 0.11</td>
<td>0.84; 0.91; 0.85</td>
</tr>
<tr>
<td>Castelporziano</td>
<td>1</td>
<td>3.31; 3.55; 3.68</td>
<td>0.61</td>
<td>−6.03; −7.08; −7.68</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.45; 1.75; 1.20</td>
<td>−2.85; −0.96; 0.08</td>
<td>0.65; 0.67; 0.80</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.69; 0.94; 1.46</td>
<td>−7.74; 0.43; −0.37</td>
<td>0.07; 0.81; 0.84</td>
</tr>
<tr>
<td>Espirra</td>
<td>1</td>
<td>4.19; 4.12; 3.61</td>
<td>−0.64</td>
<td>−6.02; −5.79; −4.23</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.85; 1.42; 1.32</td>
<td>−0.37; 0.19; 0.30</td>
<td>0.80; 0.81; 0.81</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.66; 2.01; 1.43</td>
<td>−7.71; −0.62; 0.18</td>
<td>0.35; 0.74; 0.78</td>
</tr>
</tbody>
</table>
Figure 1. A comparison of the observed (OBS) and modelled (CTRL) Latent Heat (LE), transpiration (E) and gross primary productivity (GPP) at five Fluxnet sites during 2003. The grey shading highlights the heatwave period between the 1 June and the 31 August. The data have been smoothed with a 5 day moving window to aid visualisation.
Figure 2. Modelled impact of drought on the assimilation rate ($A$), shown as (a) a function of volumetric soil moisture content ($\theta$) and (b) soil water potential ($\Psi_s$) for a sand and clay soil.
Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.

Figure 3. A comparison of the observed (OBS) and modelled latent Heat (LE), transpiration ($E$) and gross primary productivity (GPP) at the Tharandt site during 2003. Simulations show the control (CTRL) and the three drought sensitivities to drought (high, medium, low) based on Zhou et al. (2013, 2014) and three different methods to calculate soil water potential ($\Psi_s$). The grey shading highlights the heatwave period between the 1 June and the 31 August. The data have been smoothed with a 5 day moving window to aid visualisation.
Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.

Figure 4. A comparison of the observed (OBS) and modelled latent Heat (LE), transpiration (E) and gross primary productivity (GPP) at the Hesse site during 2003. Simulations show the control (CTRL) and the three drought sensitivities to drought (high, medium, low) based on Zhou et al. (2013, 2014) and three different methods to calculate soil water potential ($\Psi_S$). The grey shading highlights the heatwave period between the 1 June and the 31 August. The data have been smoothed with a 5 day moving window to aid visualisation.
Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.

---

**Figure 5.** A comparison of the observed (OBS) and modelled latent Heat (LE), transpiration (E) and gross primary productivity (GPP) at the Roccarespampani site during 2003. Simulations show the control (CTRL) and the three drought sensitivities to drought (high, medium, low) based on Zhou et al. (2013, 2014) and three different methods to calculate soil water potential ($\Psi_s$). The grey shading highlights the heatwave period between the 1 June and the 31 August. The data have been smoothed with a 5 day moving window to aid visualisation.
Do land surface models need to include differential plant species responses

M. G. De Kauwe et al.

Figure 6. A comparison of the observed (OBS) and modelled latent Heat (LE), transpiration ($E$) and gross primary productivity (GPP) at the Castelporziano Fluxnet site during 2003. Simulations show the control (CTRL) and the three drought sensitivities to drought (high, medium, low) based on Zhou et al. (2013, 2014) and three different methods to calculate soil water potential ($\Psi_s$). The grey shading highlights the heatwave period between the 1 June and the 31 August. The data have been smoothed with a 5 day moving window to aid visualisation.
Figure 7. A comparison of the observed (OBS) and modelled latent Heat (LE), transpiration ($E$) and gross primary productivity (GPP) at the Espirra site during 2003. Simulations show the control (CTRL) and the three drought sensitivities to drought (high, medium, low) based on Zhou et al. (2013, 2014) and three different methods to calculate soil water potential ($\Psi_S$). The grey shading highlights the heatwave period between the 1 June and the 31 August. The data have been smoothed with a 5 day moving window to aid visualisation.