Global vegetation gross primary production estimation using satellite-derived light-use efficiency and canopy conductance

Yebra, Marta^{1,2,3}, Van Dijk, Albert I.J.M.^{1,2,3}, Leuning, Ray² and Guerschman, Juan Pablo⁴

1 Fenner School of Environment and Society, the Australian National University, ACT,
Canberra, Australia.

2 CSIRO Oceans and Atmosphere Flagship, ACT, Canberra, Australia.

3 Bushfire & Natural Hazards Cooperative Research Centre, Melbourne, Australia.

4 CSIRO Land and Water Flagship, ACT, Australia

Abstract

Climate and physiological controls of vegetation gross primary production (GPP) vary in space and time. In many ecosystems, GPP is primary limited by absorbed photosynthetically-active radiation; in others by canopy conductance. These controls further vary in importance over daily to seasonal time scales. We propose a simple but effective conceptual model that estimates GPP as the lesser of a conductance-limited (F_c) and radiation-limited (F_r) assimilation rate. F_c is estimated from canopy conductance while F_r is estimated using a light use efficiency model. Both can be related to vegetation properties observed by optical remote sensing. The model has only two fitting parameters: maximum light use efficiency, and the minimum achieved ratio of internal to external CO₂ concentration. The two parameters were estimated using data from 16 eddy covariance flux towers for six major biomes including both energy- and water-limited ecosystems. Evaluation of model estimates with flux tower-derived GPP compared favourably to that of more complex models, for fluxes averaged; per day ($r^2=0.72$, root mean square error, RMSE= 2.48 $\mu\text{mol C m}^2 \text{ s}^{-1}$, relative percentage error, RPE=-11%), over 8-days periods ($r^2=0.78$ RMSE= 2.09 $\mu\text{mol C$

25 m^2s^{-1} , RPE=-10%), over months ($r^2=0.79$, RMSE= $1.93\mu\text{mol C m}^2 \text{ s}^{-1}$, RPE=-9%) and over
26 years ($r^2=0.54$, RMSE= $1.62 \mu\text{mol C m}^2 \text{ s}^{-1}$, RPE=-9%). Using the model we estimated global
27 GPP of 107 PgC y^{-1} for 2000–2011. This value is within the range reported by other GPP
28 models and the spatial and inter-annual patterns compare favourably. The main advantages of
29 the proposed model are its simplicity, avoiding the use of uncertain biome- or land-cover
30 class mapping, and inclusion of explicit coupling between GPP and plant transpiration.

31

32 Keywords: Photosynthesis, canopy conductance, light use efficiency, FLUXNET, MODIS,
33 gross primary production, GPP, vegetation.

34 **1. Introduction**

35 The transport of CO_2 from the atmosphere into plant leaves, where it is used in
36 photosynthesis, is inextricably linked to the simultaneous transport of water vapour in the
37 opposite direction (transpiration). Plant physiological control of these opposing fluxes is
38 exerted by stomata and the degree of control is quantified in terms of leaf stomatal
39 conductance. At the ecosystem level, canopy conductance for water vapour (G_{cw}) and CO_2
40 (G_{cc}) provide links between transpiration and photosynthesis, respectively. Estimates of
41 canopy conductance can be obtained by up-scaling stomatal conductances for all leaves in the
42 canopy (Kelliher et al. 1995), or be inferred from ecosystem level measurements of
43 exchanges of water vapour and CO_2 (Baldocchi 2008). Both approaches have been shown to
44 be suitable for application at canopy or local scales ($< 1\text{-}2 \text{ km}$) but to derive regional or
45 global estimates of canopy conductance, satellite remote sensing based methods are needed.
46 In a previous study, Yebra et al. (2013) used eddy covariance measurements of water vapour
47 fluxes at 16 sites distributed globally to establish relationships between G_{cw} and Moderate
48 Resolution Imaging Spectroradiometer (MODIS) reflectance observations. When the derived

estimates of G_{cw} were combined with net radiation, wind speed and humidity deficit data, the resulting estimates of evapotranspiration (ET) compared favourably with those from alternative approaches. Moreover, the method allowed a single parameterisation for all land cover types, which avoids artefacts resulting from errors in vegetation classification. In principle, the same satellite-derived G_{cw} values can be used within a process-based model for Gross Primary Production (GPP) while providing a direct link to the coupled energy and water balance of plant canopies.

In many ecosystems, GPP is limited by the amount of absorbed photosynthetically-active radiation (APAR), rather than by canopy conductance. The simplest approach to estimating GPP for these conditions is to multiply APAR by a light-use efficiency term (LUE or ϵ , mol C mol⁻¹ APAR) representing the plant's capacity to convert light into fixed carbon (Running et al. 1999; Sims et al. 2008; Sjöström et al. 2011). This approach requires maximum LUE to be modified where or when environmental conditions limit the rate of photosynthesis. In particular, a lack of soil water leads to stomatal closure, which reduces both ET and GPP. Over longer periods, sustained reduction in water availability will reduce vegetation cover, APAR and hence GPP (Andela et al. 2013).

In this paper the study of Yebra et al. (2013) is extended to allow the prediction of GPP globally. Our aim was to test a simple model that links GPP and ET through canopy conductance, while retaining the smallest number of 'free' fitting parameters necessary to construct a well-performing model that can be used at a global scale, without the need for ancillary information on land cover class. To account for the radiation limitation of GPP, we calibrate a simple LUE model that uses MODIS remote sensing data to estimate APAR, LUE and GPP. The lesser of the two estimates of GPP based on LUE or G_{cw} were assigned to each MODIS pixel encompassing a flux tower and globally. The results were then compared to the

official MODIS GPP product (Zhao et al. 2005) and to estimates from a regression tree approach (Jung et al. 2009).

2. Theory

We use a ‘big-leaf’ description of the plant canopy and estimate the mean GPP (symbolised by F $\mu\text{mol C m}^{-2} \text{s}^{-1}$) as the lesser of a conductance-limited and radiation-limited assimilation rate, denoted by F_c and F_r , respectively:

$$F = \min (F_c, F_r) \quad (1)$$

If it is assumed that transport of CO_2 from the bulk air to the intercellular leaf space is limited by molecular diffusion through the stomata, then F_c can be calculated from G_{cw} as:

$$F_c = c_g G_{cw} (1 - R_0) C_a \quad (2)$$

where G_{cw} (in m s^{-1}) is canopy conductance to water vapour and $R_0 = C_i / C_a$ the achieved *minimum* ratio of internal (C_i) to external (C_a) CO_2 concentration (mol mol^{-1}), and the conversion coefficient c_g ($26 \text{ mol C per m}^{-3}$) relates G_{cw} (m s^{-1}) to the conductance for CO_2 in molar units (G_{cc} , $\mu\text{mol C m}^{-2} \text{s}^{-1}$) (it can be calculated as $41.6 \text{ mol C m}^{-3}$ following the ideal gas law for standard air pressure and 25°C temperature, divided by 1.6 to account for the lesser diffusivity of CO_2 compared to H_2O). If it can be assumed that R_0 is constant for a given vegetation community or at least relatively narrowly constrained, then Eq (2) can be used to estimate the maximum rate of CO_2 uptake for a given value of G_{cw} . Support for assuming a narrow range for R_0 is given by Figure 3c in Schulze et al. (1994). They extracted data from the literature for maximum surface conductance (G_{sw}) and maximum assimilation rates (F_c) for various vegetation types across the globe. A plot of F_c versus G_{sw} (their Fig. 3C) yields a slope of 1.048 with an r^2 of 0.66. Using this value into Eq. (2) and $C_a = 360$ ppm results in $(1 - R_0) = 0.11$. The corresponding value of $R_0 = 0.89$ is for optimal conditions and is expected to be lower when various environmental factors limit photosynthesis (Tuzet

et al. 2003). Here we adopt a global value of $R_0 = 0.76$ which was obtained by fitting Eq. (2) to flux station data from 16 sites distributed globally across six biomes (see section 4 below). Radiation-limited GPP (F_r) was estimated using Eq. (3), where f_{PAR} is the fraction of absorbed PAR, Q is incident PAR ($\text{mol photons m}^{-2} \text{ s}^{-1}$) and ε is a light use efficiency ($\text{mol C mol}^{-1} \text{ photons}$).

$$F_r = \varepsilon f_{PAR} Q \quad (3)$$

Most enzyme-mediated reactions have an optimum temperature range, and several other algorithms adjust GPP estimates for T (e.g. Yuan et al. (2007)). Consequently, temperature was tested for inclusion as part of algorithm development, but rejected because model performance was only very marginally improved (see supplementary material)

3. Data

MODIS-derived reflectances and canopy conductance to water vapour

The 16-day Terra-Aqua MODIS nadir reflectance product (MCD43A4, 500 m; (Strahler et al. 1999)) provides surface reflectance corrected for the bidirectional reflectance distribution function (BRDF) and atmospheric effects, creating an apparent reflectance that is not affected by the location of the sensor relative to the pixel at the time of acquisition (Schaaf et al. 2002). Subsets of MCD43A4 data for each 500 m pixel containing a flux station were retrieved for the period 2000-2012 from the MODIS Web service (http://daac.ornl.gov/MODIS/MODIS-menu/modis_webservice.html) in order to calibrate and validate our approach. For global GPP estimates we used the 0.05° (ca. 5600 m) resolution MCD43C4 global reflectance product (collection 5) for the same period. The imagery was downloaded from the Land Processes Distributed Active Archive Center (LP DAAC, https://lpdaac.usgs.gov/data_access/data_pool) and the quality control and state flags were used to remove pixels with partial or complete cloud cover or low pixel quality in the study

areas. Global estimates of canopy conductance based on remote sensing (G_{CRS}) were calculated as described by Yebra et al. (2013) for 8-day periods and at 0.05° spatial resolution. The calculations utilized three vegetation indices derived from the MCD43C4 reflectance product: the Enhanced Vegetation Index (EVI) (Huete et al. 2002), the Normalized Difference Vegetation Index (NDVI) (Rouse et al. 1974) and a crop coefficient (K_c) estimated following Guerschman et al. (2009). The data are available via <http://www.wenfo.org/wald/>.

Flux tower observations

The GPP estimates and meteorological data used in developing the model were derived from the ‘free fair-use’ Fluxnet LaThuile dataset (Agarwal et al. 2010). Following Yebra et al. (2013) we analysed 16 sites that have at least five years of data from 2000 onwards, to coincide with the period of MODIS data availability. The flux stations were surrounded by homogeneous land cover within 1 km from the measurement tower (Table 1) to ensure that the results are not compromised if some of the MODIS pixels are not fully centred on the tower (Goerner et al. 2009). Homogeneity was assessed visually, as judged by colour and texture, using high spatial resolution aerial and satellite images from various sources (Google Earth™ <http://earth.google.com>). The selected sites are located across several continents and included six major biomes, following the International Geosphere-Biosphere Programme classification scheme (Hansen 2000): woody savannas (WSA), grasslands (GRA), croplands (CRO), evergreen needle-leaf (ENF), evergreen broadleaf (EBF) and deciduous broadleaf forest (DBF). In ecohydrological terms, both energy-limited (i.e., potential evaporation (PET) < precipitation (P)) and water-limited (PET > P) ecosystems are represented. Table 1 presents the values of a wetness index (WI), computed as the ratio between the long-term (1950-2000) annual average P and annual average PET. Sites with WI > 1 are described as energy-limited while areas with WI < 1 are termed water-limited. Here we define as

$$PET = \alpha_{PT} sR_n / (s + \gamma) \quad (4)$$

where s (Pa K^{-1}) is the slope of the saturation water vapour pressure versus temperature curve, γ is the psychrometric constant (Pa K^{-1}), R_n is absorbed net radiation (W m^{-2}) and $\alpha_{PT} = 1.26$ (Priestley and Taylor 1972).

Half-hourly GPP and meteorological data were quality-checked using the flags included in the Fluxnet La Thuile dataset. Half-hourly values were averaged over daylight hours, defined as intervals with incoming shortwave radiation $> 5 \text{ W m}^{-2}$. The total daytime shortwave radiation (R_g , W m^{-2}) measured at the flux towers was used to calculate PAR assuming that the latter is 45% of the former (Howell et al. 1983). This was converted to $\mu\text{mol photon m}^{-2} \text{ s}^{-1}$ by considering that 1 J of PAR corresponds to $\sim 4.4 \mu\text{mol photons}$.

Global meteorological data

Global estimates of daily GPP at 1° resolution were calculated using meteorological data produced using the methods as described in Sheffield et al. (2006). These data are derived through a combination of reanalysis and field data and are available from Princeton University (<http://hydrology.princeton.edu>). The data used include 24 hours mean downwelling shortwave radiation flux (R_g , W m^{-2}), specific humidity (q , $\text{m}^3 \text{ m}^{-3}$), air pressure (p , Pa) and minimum (T_{min} , K) and maximum (T_{max} , K) temperature for each day. To obtain daytime average R_g , the gridded data were divided by fraction daytime calculated using trigonometric equations. Daytime air temperature (T_a) was estimated using:

$$T_a = T_{min} + 0.75 (T_{max} - T_{min}) \quad (5)$$

Daytime vapour pressure deficit (D , Pa) was calculated as

$$D = e_{sat} - e \quad (6)$$

where e_{sat} (Pa) is saturation vapour pressure and e (Pa) the actual vapour pressure, estimated respectively as:

$$169 \quad e_{sat} = 610.8 \times \exp\left(\frac{17.27 \times T_a}{T_a + 237.3}\right) \quad (7)$$

$$170 \quad e = \frac{qp}{0.622} \quad (8)$$

171 Where 0.622 is the ratio of the molar masses of water vapour and dry air.

Table 1. Summary of the Fluxnet sites used in this study. WI, wetness index, computed as the ratio between the long-term (1950-2000) annual average precipitation and annual average potential evaporation. Climate refers to the Köppen and Geiger climate classification; Main Climates: (A) tropical rain climate, (C) warm temperate climate and (D) sub-arctic climate; Precipitation: (w) desert, (s) steppe, (f) fully humid and (s) summer dry; Temperature: (a) hot summer, (b) warm summer and (c) cool summer. IGBP stands for the International Geosphere-Biosphere Programme classification scheme. CRO: Crops; DBF: deciduous broadleaf forest; ENF and EBF: evergreen needle-leaf and broadleaf forest respectively); WSA: woody savannas; GRA: grasslands. MF: Mixed Forest; SAV: Savannas. MODIS-UMD presents the land cover type according to the Boston University's UMD classification scheme employed by MOD17.

Site Code	Latitude (°)	Longitude (°)	Köppen	Average			MODIS- UMD	Vegetation description	Reference
				Precipitation (mm y ⁻¹) *=irrigated	Temperature (°C)	WI			
AU-How	-12.49	131.15	Aw	1450	27	0.88	WSA	seasonal tropical savannah	Beringer et al. (2003)
AU-Tum	-35.66	148.15	Cfb	1159	10.7	1.13	EBF	wet temperate sclerophyll forest	Leuning et al. (2005)
FI-Hyy	61.85	24.29	Dfc	620	2.2	1.07	ENF	Scots pine forest	Suni et al. (2003)
IT-Ro1	42.41	11.93	Csa	764	15.4	0.69	DBF	young oak coppice	Rey et al. (2002)
IT-Ro2	42.39	11.92	Csa	760	15.4	0.69	DBF	mature oak coppice	Tedeschi et al. (2006)

NL-Loo	52.17	5.74	Cfb	786	9.4	1.17	ENF	MF	Spruce plantation	Dolman et al. (2002)
US-Bo1	40	-88.29	Dfa	991	11	0.91	CRO	CRO	maize-soybean (rotation)	Meyers and Hollinger (2004)
US-Ha1	42.54	-72.17	Dfb	1071	6.6	1.14	DBF	DBF	mixed deciduous forest	Urbanski et al. (2007)
US-Ho1	45.2	-68.74	Dfb	1070	5.3	1.21	ENF	MF	mixed forest	Hollinger et al. (2004)
US-MMS	39.32	-86.41	Cfa	1031	10.9	0.97	DBF	DBF	mixed deciduous forest	Schmid et al. (2000)
US-Ne1	41.16	-96.48	Dfa	790*	10	0.72	CRO	CRO	irrigated maize	Verma et al. (2005)
US-Ne2	41.16	-96.47	Dfa	789*	10	0.72	CRO	CRO	irrigated maize-soybean rotation	Verma et al. (2005)
US-Ne3	41.18	-96.44	Dfa	784	10.1	0.72	CRO	CRO	maize-soybean (rotation)	Verma et al. (2005)
US-Ton	38.43	-120.97	Csa	581	15.7	0.48	WSA	WSA	oak savannah woodland	Ma et al. (2007)
US-Var	38.41	-120.95	Csa	544	15.9	0.44	GRA	WSA	annual C3 grassland	Ma et al. (2007)
US-WCr	45.81	-90.08	Dfb	787	4	0.99	DBF	Mixed	mixed forest	Cook et al. (2004)

180 4. Methods

181 For the first part of the analysis, the daily meteorological and flux data were used to
 182 derive G_{cw} by inverting the Penman Monteith combination equation to yield:

$$184 \quad G_{cw} = \frac{\lambda E G_a}{\frac{s}{\gamma} R_n - \left(\frac{s}{\gamma} + 1\right) \lambda E + \left(\frac{\rho c_p}{\gamma}\right) D G_a} \quad (9)$$

185 where λ is the latent heat of evaporation (MJ kg⁻¹), E is the evaporation rate (kg m⁻² s⁻¹),
 186 G_a is aerodynamic conductance (m s⁻¹), s the slope of the saturation vapour pressure
 187 versus temperature curve (Pa K⁻¹), γ the psychrometric constant (Pa K⁻¹), R_n net
 188 available energy (W m⁻²), ρ air density (kg m⁻³) and c_p the specific heat constant
 189 pressure of air (J kg⁻¹ K⁻¹).

190 We recognise that G_{cw} in Eq. (9) should strictly be surface conductance (G_{sw}) which
 191 accounts for both canopy transpiration (T) and evaporation from the soil (E_s). In this
 192 study we selected data corresponding to a dry canopy and soil surface (by only
 193 including observations after two days without precipitation) with NDVI > 0.4,
 194 conditions under which $E_s \ll T$ and hence $G_{cw} \approx G_{sw}$ (see Eq. 6 in Leuning et al. (2008)).
 195 Aerodynamic conductance G_a (m s⁻¹) was calculated using

$$196 \quad G_a = \frac{1}{k^2 U} \ln \left(\frac{z-d}{z_0} \right) \ln \left(\frac{z-d}{z_{0H}} \right) \quad (10)$$

197 here k is the von Karman constant (0.40), U wind speed (m s⁻¹) at the measurement
 198 height (z) (m), d the zero-plane displacement (m) and z_0 and z_{0H} the roughness lengths
 199 for momentum and heat (m), respectively. The quantities d , z_0 and z_{0H} were estimated as
 200 $0.66h$, $0.123h$ and $0.0123h$, respectively, where h is canopy height (m).

201 The atmospheric concentration of CO₂ (C_a , mol mol⁻¹) was estimated for each day of the
 202 year (in fractional year, y) using a quadratic equation fitted to the CO₂ concentrations
 203 measured at Mauna Loa:

$$204 \quad C_a = 1.206 \cdot 10^{-8}y^2 - 4.641 \cdot 10^{-5}y + 0.045 \quad . \quad (11)$$

205 The fraction absorbed PAR radiation (f_{PAR}) was calculated from the scaled Normalised
 206 Difference Vegetation Index ($NDVI^*$) derived from the MCD43A4 reflectance data,
 207 using the ramp function proposed by Donohue et al. (2008):

208

$$209 \quad f_{PAR} = f_{PAR,max} NDVI^* \quad (12)$$

210

211 where $f_{PAR,max} = 0.95$ and

$$212 \quad NDVI^* = \max \left(\min \left(\frac{NDVI - 0.1}{0.9 - 0.1}, 1 \right), 0 \right)$$

213 The assumptions underlying this function are that (a) surfaces with a $NDVI \leq 0.1$ have
 214 no vegetation cover, (b) surfaces with $NDVI \geq 0.90$ have full canopy cover, and (c)
 215 canopy cover increases linearly with $NDVI$ for intermediate values.

216 The remaining unknown variables in Eqs. (2) and (3) are R_0 and ε . The first of these
 217 was assumed to vary between sites but be invariant in time.

218 We used remotely sensed EVI to estimate ε (cf. Drolet et al. 2008; Goerner et al. 2009;
 219 Wu et al. 2009). EVI was developed to optimize the vegetation signal with improved
 220 sensitivity in high biomass regions and improved vegetation monitoring through de-
 221 coupling of the canopy background signal and a reduction in atmosphere influences
 222 (Huete et al. 2002). It was calculated from MCD43A4 reflectances, and similar to
 223 $NDVI$, EVI was scaled (denoted EVI^*) between values assumed to represent bare
 224 surfaces ($EVI = 0.05$) and vegetation with maximum feasible light use efficiency (EVI

225 = 0.9) (A. Huete, University of Technology, Sydney, personal communication), before
 226 multiplying it with an estimated value of ε_{max} :

$$227 \quad \varepsilon = \varepsilon_{max} EVI^* \quad (13)$$

228 where

$$EVI^* = \max \left(\min \left(\frac{EVI - 0.05}{0.90 - 0.05}, 1 \right), 0 \right)$$

229 During model development, we also tested alternative remote sensing predictors of ε ,
 230 including NDVI, GVMi, unscaled EVI, fPAR and G_{CRS} (cf. Yebra et al., 2013). Each
 231 was tested in isolation and in multiplicative combinations of two indices. The
 232 formulation in Eq. (13) performed best for 12 out of 16 sites, as well as overall.
 233 Initially, we used G_{cw} values directly derived from the flux tower data by inversion of
 234 the Penman-Monteith equation (Yebra et al. 2013) and optimised a single set of values
 235 of R_0 and ε_{max} across the 16 sites (cross-site optimization). Optimal values were found
 236 by minimizing the least squared difference between observed and modelled GPP, using
 237 the Levenberg–Marquardt nonlinear optimization algorithm adapted to Interactive Data
 238 Language (IDL, ITT Visual Information Solution, Inc.) by Markwardt (2008). The code
 239 (function MPFITFUN) is available from <http://purl.com/net/mpfit>. Subsequently, a
 240 similar optimization was performed, but this time for each site individually (per-site
 241 optimization), to assess the uncertainty in the model parameterization and the influence
 242 this had on model performance. In both cases this was assessed by comparing the
 243 difference between mean modelled and flux tower GPP in relative terms (relative
 244 predictive error, $RPE = (F_{Tower} - F_{Mod})/F_{Tower}$) and the Root Mean Square Error
 245 (RMSE).

246 Furthermore, the influence of temporal scale on model performance was assessed by
247 calculating the same metrics at different time resolution (i.e., daily, weekly, monthly
248 and annually).

249 *Global GPP estimates*

250 To test model performance at global scale, we first replaced the site-based G_{cw} estimates
251 by values derived from the remote sensing product (G_{cRS}) of Yebra et al. (2013). Those
252 estimates did not consider the possible effect of D on canopy conductance, although
253 empirical evidence shows that stomatal conductance decreases with increasing D .
254 To determine if this effect needed to be accounted for, we tested whether the function
255 proposed by Lohammar et al. (1980) and Leuning (1995) [ENREF 20](#) could explain
256 residual variance in the ratio of field-based over satellite-based conductance:

$$257 \quad \frac{G_{cc}}{G_{cRS}} = \frac{C_0}{1+D/D_{50}} \quad (14)$$

258

259 where C_0 is a coefficient to correct for the possibility that G_{cRS} may be an underestimate
260 of its maximum value in saturated air (G_{cmax}), and D_{50} is the value of D at which canopy
261 conductance is half of G_{cmax} .

262 The proposed GPP model was subsequently applied at daily time steps and global scale
263 for the period 2000–2011 with spatially uniform estimates for the two parameters using
264 the 0.05° resolution MODIS canopy conductance product, and reflectance-derived
265 NVDI and EVI values, combined with daytime PAR radiation, temperature and D
266 calculated from the ‘Princeton’ meteorological dataset. These were done to examine
267 which of the global datasets introduce the greatest uncertainty in monthly-mean GPP
268 estimates

For comparison, we downloaded two other global GPP datasets; the MODIS GPP product (MOD17A2, (Zhao et al. 2005)) and the Max-Planck Institute dataset (MPI, (Jung et al. 2009)) when data for both products were available for a common period (2000–2011). The MOD17A2 product is estimated with a conceptual process model that uses MODIS FPAR/LAI product (MOD15A2v5) and NCEP/NCAR reanalysis II meteorological data, with parameterisation based on a lookup table that was applied spatially using the MODIS land cover product (MOD12Q1v4) and modified based on performance against eddy-covariance GPP observations (Zhao et al. 2005). The MPI estimates were derived using a regression tree model trained on eddy-covariance GPP estimates for 178 FLUXNET sites, with NDVI from the Advance Very High Resolution Radiometer (AVHRR) Global Inventory Modelling and Mapping Studies (GIMMS) product (until 2005) and SeaWiFS (2006 -2011) sensors and ERA-interim reanalysis meteorological data as inputs (Jung et al. 2009). Therefore, although the approaches used to construct the three datasets differ, each uses a combination of optical vegetation density remote sensing and climate reanalysis and, to different degrees, are constrained by eddy-covariance observations.

5. Results

Parameter optimization

Table 2 lists values for R_0 and ε_{max} , optimised for each site. R_0 ranged between 0.66 and 0.82 for all sites except US-Ton with $R_0 = 0.20$, the lower limit of the parameter search space. Optimal values for ε_{max} ranged between 0.025 and 0.076 molC mol⁻¹ photons. The average value of R_0 was 0.7 (0.74 excluding the anomalous R_0 value obtained for US-Ton) and the mean ε_{max} was 0.049, values close to $R_0 = 0.76$ and $\varepsilon_{max} = 0.045$, which are the overall optimised values for all sites combined.

Table 2. Quantitative measures of performance of the model for GPP at daily timescale. Predictions use canopy conductance derived from the flux tower and local meteorology. The parameters used are those derived from the cross-site optimization ($R_0=0.76$; $\varepsilon_{max}=0.045$ mol mol⁻¹) and per site (see R_0 and ε_{max} values for each site in the table 2). Allowable parameter ranges used in the optimization were 0.2-0.95 and 0.001-0.1 for R_0 and ε_{max} , respectively. RPE, relative predictive error; r^2 , coefficient of determination; RMSE, root mean square error (μmolC m² s⁻¹). p_{Fr} (%) indicate the percentage of days that F is driven by radiation ($F_r < F_c$) for each site.

Site	Days	Cross-Site Optimization			Per-site Optimization					p_{Fr} (%)
		RPE	r^2	RMSE (μmol C m ² s ⁻¹)	R_0	ε_{max} (mol mol ⁻¹)	RPE	r^2	RMSE (μmol C m ² s ⁻¹)	
AU-How	1307	7.43	0.74	1.31	0.81	0.043	-2.43	0.76	1.17	78
AU-Tum	1524	-32.49	0.52	2.65	0.72	0.064	-5.94	0.53	1.85	97
FI-Hyy	1232	-22.56	0.87	1.44	0.69	0.052	-8.68	0.87	1.20	85
IT-Ro1	1980	-2.36	0.80	1.37	0.68	0.034	-5.16	0.84	1.24	53
IT-Ro2	760	-10.36	0.79	2.19	0.71	0.053	6.00	0.79	2.05	55
NL-Loo	1488	-46.07	0.76	3.02	0.66	0.076	-11.93	0.77	1.76	90
US-Bo1	1402	-1.63	0.82	2.63	0.79	0.040	-13.43	0.82	2.52	88
US-Ha1	891	0.64	0.74	2.32	0.72	0.041	-4.26	0.75	2.29	73
US-Ho1	1055	-25.55	0.79	2.11	0.67	0.054	-6.14	0.79	1.68	58
US-MMS	945	45.83	0.80	3.75	0.79	0.025	-5.3	0.83	1.62	61
US-Ne1	788	-24.65	0.85	4.05	0.77	0.060	-5.36	0.85	3.27	95
US-Ne2	691	-12.05	0.85	3.34	0.77	0.051	-2.45	0.85	3.05	95
US-Ne3	732	-22.48	0.75	3.9	0.75	0.066	4.39	0.75	3.34	90
US-Ton	1805	-27.28	0.75	1.21	0.20	0.045	-11.79	0.75	1.03	52
US-Var	1195	8.60	0.89	1.29	0.82	0.040	-8.02	0.89	1.06	73

US-WCr	1090	20.28	0.81	3.15	0.71	0.035	-0.80	0.81	2.51	84
All	18885	-10.99	0.72	2.48	-	-	-5.43	0.82	1.95	81
Min	691	-46.07	0.52	1.21	0.20	0.025	-13.43	0.53	1.03	52
Max	1980	45.83	0.89	4.05	0.82	0.076	6.00	0.89	3.34	97
Mean	1180	-9.04	0.78	2.48	0.70	0.049	-5.08	0.79	1.98	77

301

302 *Model evaluation against flux tower data*

303 Model predictions of daily GPP using the uniform parameters ($R_0=0.76$, $\varepsilon_{max}=0.045$ mol

304 mol^{-1}), canopy conductance from the tower and local meteorology was strongly

305 correlated with flux tower GPP ($r^2=0.72$, $\text{RMSE} = 2.48 \mu\text{mol C m}^2 \text{ s}^{-1}$ (Figure 1a)).

306 There was a negative bias in modelled GPP as shown by a RPE of -10.9%.

307
 308 Figure 1. Modelled GPP (F , $\mu\text{mol C m}^{-2} \text{s}^{-1}$) plotted against flux tower GPP at different
 309 time scales. Predictions use canopy conductance derived from the tower and local
 310 meteorology. The parameters used are those derived from the cross-site optimization
 311 ($R_0=0.76$ and $\varepsilon_{max}=0.045$). The short dash lines are 1:1 lines and the solid lines are linear
 312 regression lines. Also shown are statistics for model performance: n, number of cases;
 313 RMSE, root mean square error ($\mu\text{mol C m}^{-2} \text{s}^{-1}$); RPE, relative predictive error and r^2 ,
 314 coefficient of determination.
 315

316 The model performed slightly better in reproducing 8-day and monthly average GPP
 317 values (Figure 1b and c, respectively). This was evident from a higher r^2 (0.78 and 0.79),
 318 a lower RMSE (2.09 and 1.93 $\mu\text{mol C m}^2 \text{ s}^{-1}$) and less negative bias (RPE of -9.8% and
 319 -8.6%) for 8-day and monthly values, respectively. Aggregation to annual averages
 320 reduced the errors with respect to daily values (RMSE = 1.62 $\mu\text{mol C m}^2 \text{ s}^{-1}$, RPE = -
 321 8.8%, (Figure 1d), but r^2 decreased to 0.54.

322 Looking at the result of the cross-site optimization of R_0 and ϵ_{max} for individual sites, the
 323 RMSE ranged from 1.21 to 4.05 $\mu\text{mol C m}^2 \text{ s}^{-1}$ and r^2 from 0.52 to 0.89 (Table 2). The
 324 per-site optimization described daily variations in GPP about equally well as the cross-
 325 site optimization (RMSE ranged from 1.03 to 3.34 $\mu\text{mol C m}^2 \text{ s}^{-1}$ and r^2 from 0.53 to
 326 0.89 (Table 2)). However, the use of the values of R_0 and ϵ_{max} derived from the per-site
 327 optimization considerably decreased the RMSE and RPE at AU-Tum, NL-Loo and US-
 328 MMS, although the r^2 values remained similar. Pooling data from the per-site
 329 optimization for all sites resulted in $r^2 = 0.82$, RMSE = 1.95 $\mu\text{mol C m}^2 \text{ s}^{-1}$ and RPE = -
 330 5.43%.

331 Figure 2 shows that the algorithm correctly predicted the timing of the annual cycle in
 332 GPP at each of the mid-latitude sites, largely because of the seasonal variation in
 333 incoming shortwave radiation, a key input to the model. In contrast, the seasonal
 334 amplitude of 8-day GPP is underestimated by 20 – 40% at five of the 16 sites (AU-Tum,
 335 NL-Loo, US-Ne1,2,3) and is overestimated by a similar percentage at two others (US-
 336 MMS, US-WCr).

337 The model indicates that GPP is driven by radiation ($F_r < F_c$) for > 50% of the time at
 338 each site (52 % $\langle p_{Fr} \rangle$ 97%, Table 2). However, conductance plays an almost equally
 339 important role at IT-Ro1, IT-Ro2, US-Ton and US-Ho1 ($p_{Fr} \approx 50\%$) while solar

340 radiation is the primary limiting factor at AU-Tum, NL-Loo, FI-Hyy, US-Bo1 and the
 341 Nebraska crop sites ($p_{Fr} \geq 85\%$).
 342 Radiation and canopy conductance constrained GPP at different times during the
 343 growing season. For example, GPP at US-Var was strongly limited by conductance
 344 during the hot dry summers but was limited by radiation during the rest of the year
 345 (Figure 3).

346
 347 Figure 2. Variation in predicted (red open circles) and flux tower (black filled circles) 8-
 348 day mean values of GPP (F , $\mu\text{mol C m}^{-2} \text{s}^{-1}$) at model calibration sites. Predictions use
 349 canopy conductance derived from the tower and local meteorology. The parameters
 20

350 used are those derived from the cross-site optimization ($R_0=0.76$ and $\varepsilon_{max}=0.045$ mol
 351 mol⁻¹).

352

353

354 Figure 3. Variation of 8-day mean values of GPP (F , $\mu\text{mol C m}^{-2} \text{s}^{-1}$) observed at US-
 355 Var and model estimates when GPP was conductance-limited (F_c) or radiation-limited
 356 (F_r). Predictions use canopy conductance derived from the tower and local meteorology.
 357 The parameters used are those derived from the cross-site optimization ($R_0=0.76$ and
 358 $\varepsilon_{max}=0.045$). GPP was strongly limited by conductance during the hot dry summers but
 359 limited by radiation during the rest of the year.

360

361 *Response of satellite-derived canopy conductance to vapour pressure deficit*

362 The ratios of canopy conductance derived from flux tower measurements to that based
 363 on satellite observations (G_{cw}/G_{CRS}) decreased by about 50% as D increased from 0.5 kPa
 364 to 2.5 kPa (Figure 5). This result is consistent with the hyperbolic function used by
 365 Leuning (1995) for the response of stomatal conductance to D and confirms that the
 366 influence of D on canopy conductance can be accounted for by the same function. The

values of C_0 and D_{50} in Eq. (14) were optimized across sites, producing values of $C_0 =$
 1.94 and $D_{50} = 0.70$ kPa and these values were used to compute G_{CRS} .

Figure 4. Box plots showing the distribution of daily values of the ratio of flux tower derived canopy conductance (G_{cw} , derived from the PM approach) and modelled with remote sensing data (G_{CRS}) for different D intervals. Whiskers show the outer 5% percentiles.

Performance of global GPP estimates

Table 3 shows that replacing G_{cw} with G_{CRS} reduced r^2 by 7%, and increased the RMSE by 33%, while decreasing the magnitude of RPE. Replacing local with global meteorological data did not degrade the model performance appreciably. The greatest degradation in model performance was found when 0.05° reflectance data were used, although we consider model performance to be acceptable ($r^2 = 0.61$, RMSE = $2.65 \mu\text{mol C m}^{-2} \text{ s}^{-1}$, RPE = -15%).

381 Despite its simplicity, the global version of our model performed similar to MPI and
 382 MOD17 (Table 3). At monthly time scale, it produced the lowest RPE and performed
 383 better than MOD17 overall, although MPI achieved somewhat better r^2 and RMSE.

384

385 Table 3. Quantitative measures of model performance at monthly time scale when local data
 386 are replaced by estimates from global datasets. The parameters used are those derived from the
 387 cross-site optimization ($R_0=0.76$; $\epsilon_{max}=0.045 \text{ mol mol}^{-1}$ for the photosynthesis model and
 388 $D_{50}=700$ and $C_1=1.94$ for the canopy conductance model). The performance of the MPI and
 389 MODIS algorithms for the same observations is included for comparison. RPE, relative
 390 predictive error; r^2 , coefficient of determination; RMSE, root mean square error.

391

<i>Model</i>	<i>RPE (%)</i>	<i>r²</i>	<i>RMSE ($\mu\text{mol C m}^{-2} \text{ s}^{-1}$)</i>
This study (site data)	-8.55	0.79	1.93
This study (with global G_{CRS})	-3.99	0.72	2.26
This study (with global meteorology)	-7.01	0.79	1.90
This study (with global G_{CRS} and meteorology)	-3.57	0.70	2.30
This study (with global G_{CRS} , meteo and reflectances)	-15.46	0.61	2.65
MPI	-17.95	0.73	2.36
MOD17	-27.57	0.50	3.15

392
393 Figure 5. Annual average gross primary production GPP ($\text{gC m}^{-2} \text{y}^{-1}$) estimates for the
394 period 2000-2013
395

396

397 Figure 6. Annual average terrestrial GPP for 2000–2011 calculated (a) for different
 398 latitudes (in gC m⁻² y⁻¹; zero GPP was assigned to water bodies) (b) per generalised
 399 biome category (PgC y⁻¹), using the MODIS land cover product (low latitudes are
 400 defined as within ±60°).

401

402

403 Application of the GPP algorithm with constant values for R_0 and ε_{max} gave an annual
 404 average global GPP of 107 PgC y⁻¹ for 2000–2011. Global estimates for the same period
 405 derived from MOD17 and MPI were 112 (+4%) and 122 (+14 %) PgC y⁻¹, respectively.
 406 The global distribution of GPP shown in Figure 5 is very similar to that in the MOD17
 407 and MPI datasets (compare with Fig. 1B in (Beer et al. 2010)). This agreement in
 408 distribution pattern is also seen in comparisons by latitude and by biome (Figure 6). The
 409 latitudinal pattern in GPP is very similar for all three datasets but the magnitude of our
 410 estimates and of those from MOD17 are lower than those from MPI around 30 °N and
 411 15 °S (Figure 6a). Expressed per biome, our estimates were between those from the
 412 MOD17 and MPI for 4 out of 9 biomes (Figure 6b). However, our GPP estimates are
 413 lower for water-limited and sparsely vegetated ecosystems (low-latitude shrubland,
 414 savannas and ‘other’) and needleleaf forests.
 415 Anomalies in annual global GPP relative to mean values for 2000-2011 are compared in
 416 Figure 7 for the three datasets. Our algorithm yielded an inter-annual standard deviation
 417 (± 0.65 PgC y⁻¹) that is 30% lower than that from MOD17 (0.90 PgC y⁻¹), but both were
 418 considerably less than that from the MPI dataset (± 1.67 PgC y⁻¹). There was some
 419 agreement in temporal anomaly patterns between our estimates and those from MOD17
 420 ($r^2=0.34$, RMSE 0.71 PgC y⁻¹), but not with those from MPI ($r^2=0.17$, RMSE 1.46 PgC
 421 y⁻¹). Analysis per biome revealed that savannas and grasslands were the main source of
 422 inter-annual GPP variation in each of the datasets, and were responsible for most of the
 423 agreement with MOD17 ($r^2=0.81$) but also for much of the remaining disagreement
 424 (RMSE=0.35 PgC y⁻¹). The sources of temporal disagreement between our estimates
 425 and those from MPI data were distributed across biomes.
 426

427 Figure 7. Predicted annual global terrestrial GPP anomalies for 2000–2011 compared
 428 with values from the MODIS and MPI GPP datasets. Anomaly is defined as the
 429 difference between annual GPP and the GPP mean for 2000–11 (107 PgC y⁻¹).
 430

431 6. Discussion

432 *Site-to-site variation in model parameters*

433 The average value for the two model-parameters R_0 and ε_{max} derived by model
 434 optimisation for each site was close to the overall optimised values for the 16 sites
 435 combined. However the analysis also showed site variability in the parameter values.
 436 NL-Loo obtained and optimized value of $\varepsilon_{max} = 0.076$ and $R_0 = 0.66$ while for US-
 437 MMS values of $R_0 = 0.79$ and $\varepsilon_{max} = 0.025$ were found. Therefore, the use of a single
 438 optimised intermediate value of ε_{max} and R_0 explains the overestimation of GPP for US-
 439 MMS and underestimation for NL-Loo.

440 Following Leuning (1995), our model predicts that there is a minimum achievable C_i/C_a
 441 ratio, R_0 , that determines maximum possible GPP for a given stomatal conductance. To
 442 test this, apparent C_i/C_a ratios can also be calculated directly from the observations, by
 443 rearranging Eq. (2) and inserting the measured GPP and G_{cw} that are derived from the
 444 flux tower measurements:

$$445 \quad R_0 = 1 - \frac{F_w}{c_g G_{cw} C_a} \quad (15)$$

446 The main uncertainty in this approach is to determine on what days GPP was in fact
 447 limited by conductance, that is, when $F=F_w$, as required by Eq. (15). It can be assumed
 448 that VPD on such days was typically high.

449

450

451 Figure 8. Relationship between vapour pressure deficit (VPD, Pa) and the ratio of
 452 internal and external CO₂ concentration (C_i/C_a) calculated from flux tower observations
 453 at 16 sites. Each point represents at least 10 days of data. Sites are coloured by IGBP

454 vegetation type (blue=ENF, brown=DBF/EBF green=CRO/GRA, black=WSA; see
455 caption Table 1 for meaning of codes).

456

457 The relationships found in Figure 8 confirm the conceptual model, and agree with the
458 general relationship between D and apparent C_i/C_a predicted by Leuning (1995). Values
459 of C_i/C_a reach a minimum (R_0) around $D > 1.5$ kPa, although not all sites experience high
460 D values, making it difficult to determine R_0 for some of the sites The relationship C_a
461 varies between sites and suggests a dependence on vegetation type: needle-leaf forests
462 and drought-adapted blue oaks (US-Ton) appear to have the lowest R_0 values, broadleaf
463 forests have intermediate values, and crops and the (savannah) grasses (AU-How and
464 US-Var) the highest values.

465

466 *Model evaluation*

467 There is good agreement between modelled and measured GPP at the 16 globally
468 distributed flux stations at daily to yearly time scales. The highest errors were found for
469 the Nebraska crop sites (US-Ne1, US-Ne2 and US-Ne3) and US-MMS although the r^2
470 values for these sites were still high (0.75-0.85). At US-Ne1 (maize), the peak in GPP
471 was always underestimated whereas for US-Ne2 and US-Ne3 (maize-soybean rotation
472 sites) the peak was underestimated only for years with a maize crop (2001 and 2003)
473 and not in seasons with a soybean crop. Maize assimilates CO_2 via a C4 fixation
474 pathway. It is well-established that C4 species photosynthesise more effectively than C3
475 species and reach a maximum rate at lower radiation level. Consequently different ϵ_{max}
476 values might be required for ecosystems dominated by C3 or C4 species. Still et al.
477 (2003) developed an approach for capturing the spatial and temporal heterogeneity of

478 both photosynthetic types by combining remote sensing products, physiological
 479 modelling, a spatial distribution of global crop fractions, and national harvest area data
 480 for major crop types. The derived information on the proportion of C3 and C4 plants
 481 per pixel used in such an approach could then be used in our algorithm (Donohue et al.
 482 2014; Ryu et al. 2011). The main disadvantage is that this increases the complexity of
 483 the model and potentially degrades model performance due to considerable uncertainty
 484 in C3/C4 mapping.

485 Overall, our model captures the main factors that constrain photosynthesis. Radiation-
 486 limited photosynthesis dominated overall, but conductance played an almost equally
 487 limiting role at sites in Mediterranean climates with a lower wetness index ($WI < 0.69$).
 488 However, the influence of conductance limitation was also strong at US-Ho1 although
 489 this is a humid site ($WI = 1.21$). GPP was mainly driven by solar radiation at other
 490 temperate, continental and crop sites where water availability was usually adequate (WI
 491 > 0.9), confirming earlier studies (van Dijk et al. 2005; van Gorsel et al. 2013).

492 Our model compares favourably against the accuracy of MPI and MOD17 as well as
 493 other published models. For example, the model of Yuan et al. (2010) explained about
 494 75% and 61% of the variation of 8-day GPP estimates at calibration and validation sites,
 495 respectively, whereas our model explained 78% of 8-day tower GPP at calibration sites.
 496 Compared with the r^2 values reported by Yuan et al. (2010) (their Table 2) our model
 497 performed better at 6 sites (US-Bo, US-N1,2,3, US-WCr and US-Var, by 1-15%) and
 498 worse at two (US-Ton and US-Ho, by 5% and 9%, respectively). To compare our
 499 results with those of Ryu et al. (2011), we calculate the r^2 and RMSE modelled versus
 500 measured GPP for 7 sites common to both studies, for the year used in their analysis.

501 Our model performed similarly at FI-Hyy, US-MMS and US-WCr, less well at US-Ne1
 502 ($r^2 = 0.90$ and 0.75 , $RMSE = 2$ and $4 \text{ gC m}^{-2} \text{ d}^{-1}$ for Ryu et al. (2011) and this study,

503 respectively) and better at AU-Tum (same r^2 but RMSE was $1.8 \text{ gC m}^{-2} \text{ d}^{-1}$ lower), AU-
504 How ($r^2 = 0.30$ and 0.82 , RMSE = 2.2 and $0.9 \text{ gC m}^{-2} \text{ d}^{-1}$ for Ryu et al. (2011) and this
505 study, respectively), and AU-Ton ($r^2 = 0.30$ and 0.71 , RMSE = 1 and $1.5 \text{ gC m}^{-2} \text{ d}^{-1}$ for
506 Ryu et al. (2011) and this study, respectively).

507 *Global estimates of GPP*

508 Our analysis showed that using single, optimized parameter values ($R_0=0.76$; $\varepsilon_{max}=0.045$
509 mol mol^{-1}) in our algorithm did not degrade model performance much compared to site-
510 specific calibrations. This is perhaps surprising considering the site-to-site variability in
511 locally optimized values, and may indicate that the model is more sensitive to model
512 inputs than to the parameters. Additionally, using MODIS-derived canopy conductance
513 (Yebra et al. 2013) did not strongly degrade model performance compared to using
514 conductance derived from flux measurements. We extended the algorithm of Yebra et
515 al. (2013) to account for the response of canopy conductance to D . The benefit of our
516 conductance-based approach is that it produces mutually consistent estimates of GPP
517 and transpiration, and therefore can be used to estimate canopy-level water use
518 efficiency.

519 Further degradation was expected when global meteorological forcing data were used
520 rather than site observations but the differences in model performance were small
521 (Table 3). Overall the differences between gridded temperature and radiation estimates
522 produced by Sheffield et al. (2006) and observations from the tower were modest
523 (results not shown). However, it is noted that most of the 16 sites are in regions with a
524 relatively good measurement network. Errors are likely to be larger in some other parts
525 of the world where measurements networks are less developed.

526 A second important advantage of the approach developed here is that identical
 527 coefficients can be applied globally to all land cover types. Although this increased
 528 RMSE slightly at some sites, when compared to using site specific parameters, it avoids
 529 errors induced by categorical land cover mapping. Table 1 shows that classification of
 530 half of the 16 sites differed between the Boston University's UMD classification
 531 scheme employed by MOD17, and the IGBP classification scheme. This
 532 misclassification affects the MODIS GPP estimates through the set of parameters that
 533 are applied to each pixel. Furthermore, our approach does not depend on indirect
 534 satellite-derived vegetation products such as LAI, which have been reported to increase
 535 errors in GPP estimation. For example, Ryu et al. (2011) found substantial
 536 overestimations of GPP at the AU-Tum site due to inaccuracies in MODIS LAI. They
 537 also found that underestimations of LAI during spring or autumn caused the
 538 underestimation of their model derived GPP in most needle-leaf forests.
 539 Our global-average GPP estimates for 2000–2011 were 4% lower than those from
 540 MOD17 and 14% lower than those from MPI. Beer et al. (2010) estimated the range of
 541 plausible GPP estimates following the MPI machine learning method at 105-125 PgC y⁻¹
 542 ¹; our estimate of 107 PgC y⁻¹ lies within that range. Furthermore, Ryu et al. (2011)
 543 estimated average GPP at 118 ± 26 PgC yr⁻¹ for 2001-2003 and Yuan et al. (2010)
 544 reported a value of 111 ± 21 PgC yr⁻¹ for 2000-2003; close to our values (also 107 PgC
 545 y⁻¹ for both periods).
 546 When assessed by biome category, our GPP estimates were similar to MOD17 and MPI
 547 estimates for temperate and tropical humid biomes, but lower for needle-leaf forests
 548 and the most severely water-limited ecosystems.
 549 The inter-annual pattern of global GPP predicted by our model is comparable to those in
 550 the MOD17 data, but both are considerably different from those in the MPI data.

551 [ENREF 52](#) Zhao and Running (2010) showed that the MOD17 pattern agrees well
552 with inter-annual variations in atmospheric CO₂ concentrations. In particular, the
553 negative anomalies in both MOD17 and our own estimates (Figure 7) could possibly
554 explain, at least in part, accelerated CO₂ increases in the atmosphere in 2002 and 2005.
555 On this basis, we conclude that the inter-annual GPP pattern derived from MPI, which
556 has a strong positive anomaly in 2005, is less plausible. We speculate that data
557 inconsistencies in the changeover from AVHRR GIMMS (up to 2005) to SeaWiFS
558 (2006 onwards) may be a factor; annual anomalies appear to agree better during the
559 latter period (Figure 7).

560 **6. Conclusions**

561 The simple algorithm we propose for the estimation of GPP performs well at daily,
562 monthly and annual time-scales and similarly well or better than other approaches
563 published in the literature. The model was applied globally to compute global estimates
564 of GPP using global meteorological data and MODIS nadir reflectances. The global
565 spatial and temporal patterns in our GPP estimates compared favourably with other
566 datasets. We consider this an encouraging result, given (a) the simplicity of our two-
567 parameter model, (b) the lack of biome- or land-cover specific parameters, and (c) the
568 simple but explicit coupling between ET and GPP. The site-based analysis and global
569 comparison suggests that perhaps the greatest simplification made was the assumption
570 that a minimum C_i/C_a ratio (i.e., R_0) of 0.75–0.85 is achievable in all ecosystems. We
571 found some evidence that lower ratios are achieved by needle-leaf and drought-tolerant
572 species, and higher ratios by crops and grasses (Table 2). Future research may help to
573 find ways to better predict R_0 while still avoiding the use of uncertain categorical land
574 cover maps.

576 **Acknowledgements**

577 This work used eddy covariance data acquired by the FLUXNET community and in
578 particular by the following networks: AmeriFlux (U.S. Department of Energy,
579 Biological and Environmental Research, Terrestrial Carbon Program
580 (DE-FG02-04ER63917 and DE-FG02-04ER63911)), CarboEuropeIP, CarboItaly and
581 OzFlux. We acknowledge the financial support to the eddy covariance data
582 harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck
583 Institute for Biogeochemistry, National Science Foundation, University of Tuscia,
584 Université Laval and Environment Canada and US Department of Energy and the
585 database development and technical support from Berkeley Water Center, Lawrence
586 Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National
587 Laboratory, University of California - Berkeley, University of Virginia. The support of
588 the Commonwealth of Australia through the Cooperative Research Centre program is
589 acknowledged. We also acknowledge the contributions by Peter Hairsine, David
590 Summers and four anonymous reviewers.

591

592 **References**

- 593 Agarwal, D.A., Humphrey, M., Beekwilder, N.F., Jackson, K.R., Goode, M.M., & van
594 Ingen, C. (2010). A data-centered collaboration portal to support global carbon-
595 flux analysis. *Concurrency and Computation: Practice and Experience*, 22, 2323-
596 2334
- 597 Andela, N., Liu, Y.Y., van Dijk, A.I.J.M., de Jeu, R.A.M., & McVicar, T.R. (2013).
598 Global changes in dryland vegetation dynamics (1988-2008) assessed by satellite
599 remote sensing: comparing a new passive microwave vegetation density record
600 with reflective greenness data. *Biogeosciences*, 10, 6657-6676
- 601 Baldocchi, D. (2008). Breathing of the terrestrial biosphere: lessons learned from a
602 global network of carbon dioxide flux measurement systems. *Australian Journal*
603 *of Botany*, 56, 1-26
- 604 Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck,
605 C., Arain, M., Baldocchi, D., Bonan, G., Bondeau, A., Cescatti, A., Lasslop, G.,
606 Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K., Rouspard, O.,
607 Veenendaal, E., Viovy, N., Williams, C., Woodward, F., & Papale, D. (2010).
608 Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation
609 with Climate. *Science*, 329, 834-838
- 610 Beringer, J., Hutley, L.B., Tapper, N.J., Coutts, A., Kerley, A., & O'Grady, A.P. (2003).
611 Fire impacts on surface heat, moisture and carbon fluxes from a tropical savanna
612 in northern Australia. *International Journal of Wildland Fire*, 12, 333-340
- 613 Cook, B.D., Davis, K.J., Wang, W.G., Desai, A., Berger, B.W., Teclaw, R.M., Martin,
614 J.G., Bolstad, P.V., Bakwin, P.S., Yi, C.X., & Heilman, W. (2004). Carbon
615 exchange and venting anomalies in an upland deciduous forest in northern
616 Wisconsin, USA. *Agricultural and Forest Meteorology*, 126, 271-295
- 617 Dolman, A.J., Moors, E.J., & Elbers, J.A. (2002). The carbon uptake of a mid latitude
618 pine forest growing on sandy soil. *Agricultural and Forest Meteorology*, 111,
619 157-170
- 620 Donohue, R.J., Hume, I.H., Roderick, M.L., McVicar, T.R., Beringer, J., Hutley, L.B.,
621 Gallant, J.C., Austin, J.M., van Gorsel, E., Cleverly, J.R., Meyer, W.S., & Arndt,
622 S.K. (2014). Evaluation of the remote-sensing-based DIFFUSE model for
623 estimating photosynthesis of vegetation. *Remote Sensing of Environment*, 155,
624 349-365
- 625 Donohue, R.J., Roderick, M.L., & McVicar, T.R. (2008). Deriving consistent long-term
626 vegetation information from AVHRR reflectance data using a cover-triangle-
627 based framework. *Remote Sensing of Environment*, 112, 2938-2949
- 628 Drolet, G.G., Middleton, E.M., Huemmrich, K.F., Hall, F.G., Amiro, B.D., Barr, A.G.,
629 Black, T.A., McCaughey, J.H., & Margolis, H.A. (2008). Regional mapping of
630 gross light-use efficiency using MODIS spectral indices. *Remote Sensing of*
631 *Environment*, 112, 3064-3078
- 632 Goerner, A., Reichstein, M., & Rambal, S. (2009). Tracking seasonal drought effects on
633 ecosystem light use efficiency with satellite-based PRI in a Mediterranean forest.
634 *Remote Sensing of Environment*, 113, 1101-1111
- 635 Guerschman, J.P., Van Dijk, A.I.J.M., Mattersdorf, G., Beringer, J., Hutley, L.B.,
636 Leuning, R., Pipunic, R.C., & Sherman, B.S. (2009). Scaling of potential
637 evapotranspiration with MODIS data reproduces flux observations and catchment
638 water balance observations across Australia. *Journal of Hydrology*, 369, 107-119

639 Hansen, M.C. (2000). A comparison of the IGBP DISCover and University of Maryland
640 1km global land cover products. *International Journal of Remote Sensing*, 21,
641 1365-1373

642 Hollinger, D.Y., Aber, J., Dail, B., Davidson, E.A., Goltz, S.M., Hughes, H., Leclerc,
643 M.Y., Lee, J.T., Richardson, A.D., Rodrigues, C., Scott, N.A., Achuatavarier, D.,
644 & Walsh, J. (2004). Spatial and temporal variability in forest-atmosphere CO₂
645 exchange. *Global Change Biology*, 10, 1689-1706

646 Howell, T.A., Meek, D.W., & Hatfield, J.L. (1983). Relationship of Photosynthetically
647 Active Radiation to Shortwave Radiation in the San-Joaquin Valley. *Agricultural
648 Meteorology*, 28, 157-175

649 Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002).
650 Overview of the radiometric and biophysical performance of the MODIS
651 vegetation indices. *Remote Sensing of Environment*, 83, 195-213

652 Jung, M., Reichstein, M., & Bondeau, A. (2009). Towards global empirical upscaling of
653 FLUXNET eddy covariance observations: validation of a model tree ensemble
654 approach using a biosphere model. *Biogeosciences*, 6, 2001-2013

655 Kelliher, F.M., Leuning, R., Raupach, M.R., & Schulze, E.D. (1995). Maximum
656 conductances for evaporation from global vegetation types. *Agricultural and
657 Forest Meteorology*, 73, 1-16

658 Leuning, R. (1995). A Critical-Appraisal of a Combined Stomatal-Photosynthesis
659 Model for C-3 Plants. *Plant Cell and Environment*, 18, 339-355

660 Leuning, R., Cleugh, H.A., Zegelin, S.J., & Hughes, D. (2005). Carbon and water fluxes
661 over a temperate Eucalyptus forest and a tropical wet/dry savanna in Australia:
662 measurements and comparison with MODIS remote sensing estimates.
663 *Agricultural and Forest Meteorology*, 129, 151-173

664 Leuning, R., Zhang, Y.Q., Rajaud, A., Cleugh, H., & Tu, K. (2008). A simple surface
665 conductance model to estimate regional evaporation using MODIS leaf area index
666 and the Penman-Monteith equation. *Water Resources Research*, 44, W10419

667 Lohammar, T., Larsson, S., Linder, S., & Falk, S.O. (1980). FAST: Simulation Models
668 of Gaseous Exchange in Scots Pine. *Ecological Bulletins*, 505-523

669 Ma, S., Baldocchi, D.D., Xu, L., & Hehn, T. (2007). Inter-annual variability in carbon
670 dioxide exchange of an oak/grass savanna and open grassland in California.
671 *Agricultural and Forest Meteorology*, 147, 157-171

672 Markwardt, C.B. (2008). Non-Linear Least Squares Fitting in IDL with MPFIT. In.
673 Quebec, Canada: Astronomical Data Analysis Software and Systems XVIII

674 Meyers, T.P., & Hollinger, S.E. (2004). An assessment of storage terms in the surface
675 energy balance of maize and soybean. *Agricultural and Forest Meteorology*, 125,
676 105-115

677 Priestley, C.H.B., & Taylor, R.J. (1972). On the Assessment of Surface Heat Flux and
678 Evaporation Using Large-Scale Parameters. *Monthly Weather Review*, 100, 81-92

679 Rey, A., Pegoraro, E., Tedeschi, V., De Parri, I., Jarvis, P.G., & Valentini, R. (2002).
680 Annual variation in soil respiration and its components in a coppice oak forest in
681 Central Italy. *Global Change Biology*, 8, 851-866

682 Rouse, J.W., Haas, R.W., Schell, J.A., Deering, D.H., & Harlan, J.C. (1974).
683 Monitoring the vernal advancement and retrogradation (Greenwave effect) of
684 natural vegetation. In. Greenbelt, MD. USA: NASA/GSFC

685 Running, S., Nemani, R., Glassy, J., & Thornton, P. (1999). MODIS daily
686 photosynthesis (PSN) and annual Net Primary Production (NPP) product
687 (MOD17). Algorithm Theoretical Basis Document. In V. 3.0 (Ed.)

688 Ryu, Y., Baldocchi, D.D., Kobayashi, H., van Ingen, C., Li, J., Black, T.A., Beringer, J.,
 689 van Gorsel, E., Knohl, A., Law, B.E., & Rouspard, O. (2011). Integration of
 690 MODIS land and atmosphere products with a coupled-process model to estimate
 691 gross primary productivity and evapotranspiration from 1 km to global scales.
 692 *Global Biogeochemical Cycles*, 25, GB4017
 693 Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C.,
 694 Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M.,
 695 Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B., Liang, S.,
 696 Privette, J.L., & Roy, D. (2002). First operational BRDF, albedo nadir reflectance
 697 products from MODIS. *Remote Sensing of Environment*, 83, 135-148
 698 Schmid, H.P., Grimmond, C.S.B., Cropley, F., Offerle, B., & Su, H.-B. (2000).
 699 Measurements of CO₂ and energy fluxes over a mixed hardwood forest in the
 700 mid-western United States. *Agricultural and Forest Meteorology*, 103, 357-374
 701 Schulze, E., Kelliher, F.M., Korner, C., Lloyd, J., & Leuning, R. (1994). Relationships
 702 among Maximum Stomatal Conductance, Ecosystem Surface Conductance,
 703 Carbon Assimilation Rate, and Plant Nitrogen Nutrition: A Global Ecology
 704 Scaling Exercise. *Annual Review of Ecology and Systematics*, 25, 629-662
 705 Sheffield, J., Goteti, G., & Wood, E.F. (2006). Development of a 50-Year High-
 706 Resolution Global Dataset of Meteorological Forcings for Land Surface
 707 Modeling. *Journal of Climate*, 19, 3088-3111
 708 Sims, D.A., Rahman, A.F., Cordova, V.D., El-Masri, B.Z., Baldocchi, D.D., Bolstad,
 709 P.V., Flanagan, L.B., Goldstein, A.H., Hollinger, D.Y., Misson, L., Monson, R.K.,
 710 Oechel, W.C., Schmid, H.P., Wofsy, S.C., & Xu, L. (2008). A new model of gross
 711 primary productivity for North American ecosystems based solely on the
 712 enhanced vegetation index and land surface temperature from MODIS. *Remote*
 713 *Sensing of Environment*, 112, 1633-1646
 714 Sjöström, M., Ardö, J., Arneth, A., Boulain, N., Cappelaere, B., Eklundh, L., de
 715 Grandcourt, A., Kutsch, W.L., Merbold, L., Nouvellon, Y., Scholes, R.J.,
 716 Schubert, P., Seaquist, J., & Veenendaal, E.M. (2011). Exploring the potential of
 717 MODIS EVI for modeling gross primary production across African ecosystems.
 718 *Remote Sensing of Environment*, 115, 1081-1089
 719 Still, C.J., Berry, J.A., Collatz, G.J., & DeFries, R.S. (2003). Global distribution of C3
 720 and C4 vegetation: Carbon cycle implications. *Global Biogeochemical Cycles*, 17,
 721 1006
 722 Strahler, A.H., Muller, J.P., & Modis Sciences Team Members (1999). MODIS BRDF
 723 Albedo Product: Algorithm Theoretical Basis Document Version 5.0. In (p. 53)
 724 Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Dal Maso, M.,
 725 Kulmala, M., & Vesala, T. (2003). Long-term measurements of surface fluxes
 726 above a Scots pine forest in Hyytiala, southern Finland, 1996-2001. *Boreal*
 727 *Environment Research*, 8, 287-301
 728 Tedeschi, V., Rey, A.N.A., Manca, G., Valentini, R., Jarvis, P.G., & Borghetti, M.
 729 (2006). Soil respiration in a Mediterranean oak forest at different developmental
 730 stages after coppicing. *Global Change Biology*, 12, 110-121
 731 Tuzet, A., Perrier, A., & Leuning, R. (2003). A coupled model of stomatal conductance,
 732 photosynthesis and transpiration. *Plant, Cell & Environment*, 26, 1097-1116
 733 Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K.,
 734 Fitzjarrald, D., Czikowsky, M., & Munger, J.W. (2007). Factors controlling CO₂
 735 exchange on timescales from hourly to decadal at Harvard Forest. *J. Geophys.*
 736 *Res.*, 112, G02020

737 van Dijk, A.I.J.M., Dolman, A.J., & Schulze, E.D. (2005). Radiation, temperature, and
 738 leaf area explain ecosystem carbon fluxes in boreal and temperate European
 739 forests. *Global Biogeochemical Cycles*, 19
 740 van Gorsel, E., Berni, J.A.J., Briggs, P., Cabello-Leblic, A., Chasmer, L., Cleugh, H.A.,
 741 Hacker, J., Hantson, S., Haverd, V., Hughes, D., Hopkinson, C., Keith, H., Kljun,
 742 N., Leuning, R., Yebra, M., & Zegelin, S. (2013). Primary and secondary effects
 743 of climate variability on net ecosystem carbon exchange in an evergreen
 744 Eucalyptus forest. *Agricultural and Forest Meteorology*, 182, 248-256
 745 Verma, S.B., Dobermann, A., Cassman, K.G., Walters, D.T., Knops, J.M., Arkebauer,
 746 T.J., Suyker, A.E., Burba, G.G., Amos, B., Yang, H.S., Ginting, D., Hubbard,
 747 K.G., Gitelson, A.A., & Walter-Shea, E.A. (2005). Annual carbon dioxide
 748 exchange in irrigated and rainfed maize-based agroecosystems. *Agricultural and*
 749 *Forest Meteorology*, 131, 77-96
 750 Wu, C., Niu, Z., Tang, Q., Huang, W., Rivard, B., & Feng, J. (2009). Remote estimation
 751 of gross primary production in wheat using chlorophyll-related vegetation indices.
 752 *Agricultural and Forest Meteorology*, 149, 1015-1021
 753 Yebra, M., Van Dijk, A., Leuning, R., Huete, A., & Guerschman, J.P. (2013).
 754 Evaluation of optical remote sensing to estimate actual evapotranspiration and
 755 canopy conductance. *Remote Sensing of Environment*, 129, 250-261
 756 Yuan, W., Liu, S., Yu, G., Bonnefond, J.-M., Chen, J., Davis, K., Desai, A.R.,
 757 Goldstein, A.H., Gianelle, D., Rossi, F., Suyker, A.E., & Verma, S.B. (2010).
 758 Global estimates of evapotranspiration and gross primary production based on
 759 MODIS and global meteorology data. *Remote Sensing of Environment*, 114, 1416-
 760 1431
 761 Yuan, W.P., Liu, S., Zhou, G.S., Zhou, G.Y., Tieszen, L.L., Baldocchi, D., Bernhofer,
 762 C., Gholz, H., Goldstein, A.H., Goulden, M.L., Hollinger, D.Y., Hu, Y., Law,
 763 B.E., Stoy, P.C., Vesala, T., & Wofsy, S.C. (2007). Deriving a light use efficiency
 764 model from eddy covariance flux data for predicting daily gross primary
 765 production across biomes. *Agricultural and Forest Meteorology*, 143, 189-207
 766 Zhao, M., Heinsch, F.A., Nemani, R.R., & Running, S.W. (2005). Improvements of the
 767 MODIS terrestrial gross and net primary production global data set. *Remote*
 768 *Sensing of Environment*, 95, 164-176
 769 Zhao, M.S., & Running, S.W. (2010). Drought-Induced Reduction in Global Terrestrial
 770 Net Primary Production from 2000 Through 2009. *Science*, 329, 940-943
 771
 772
 773

774 **Figures**

775
776 Figure 1. Modelled GPP (F , $\mu\text{mol C m}^2 \text{ s}^{-1}$) plotted against flux tower GPP at different
777 time scales. Predictions use canopy conductance derived from the tower and local
778 meteorology. The parameters used are those derived from the cross-site optimization
779 ($R_0=0.76$ and $\varepsilon_{max}=0.045$). The short dash lines are 1:1 lines and the solid lines are linear
780 regression lines. Also shown are statistics for model performance: n , number of cases;
781 RMSE, root mean square error ($\mu\text{mol C m}^2 \text{ s}^{-1}$); RPE, relative predictive error and r^2 ,
782 coefficient of determination.

783
784 Figure 2. Variation in predicted (red open circles) and flux tower (black filled circles) 8-
785 day mean values of GPP (F , $\mu\text{mol C m}^2 \text{ s}^{-1}$) at model calibration sites. Predictions use
786 canopy conductance derived from the tower and local meteorology. The parameters
787 used are those derived from the cross-site optimization ($R_0=0.76$ and $\varepsilon_{max}=0.045 \text{ mol}$
788 mol^{-1}).

789
790 Figure 3. Variation of 8-day mean values of GPP (F , $\mu\text{mol C m}^2 \text{ s}^{-1}$) observed at US-
791 Var and model estimates when GPP was conductance-limited (F_c) or radiation-limited
792 (F_r). Predictions use canopy conductance derived from the tower and local meteorology.
793 The parameters used are those derived from the cross-site optimization ($R_0=0.76$ and
794 $\varepsilon_{max}=0.045$). GPP was strongly limited by conductance during the hot dry summers but
795 limited by radiation during the rest of the year.

796
797 Figure 4. Box plots showing the distribution of daily values of the ratio of flux tower
798 derived canopy conductance (G_{cw} , derived from the PM approach) and modelled with

799 remote sensing data (G_{CRS}) for different D intervals. Whiskers show the outer 5%
800 percentiles.

801

802 Figure 5. Annual average gross primary production GPP ($\text{gC m}^{-2} \text{y}^{-1}$) estimates for the
803 period 2000-2013

804

805 Figure 6. Annual average terrestrial GPP for 2000–2011 calculated (a) for different
806 latitudes (in $\text{gC m}^{-2} \text{y}^{-1}$; zero GPP was assigned to water bodies) (b) per generalised
807 biome category (PgC y^{-1}), using the MODIS land cover product (low latitudes are
808 defined as within $\pm 60^\circ$).

809

810 Figure 7. Predicted annual global terrestrial GPP anomalies for 2000–2011 compared
811 with values from the MODIS and MPI GPP datasets. Anomaly is defined as the
812 difference between annual GPP and the GPP mean for 2000-11 (107 PgC y^{-1}).

813

814 Figure 8. Relationship between vapour pressure deficit (VPD, Pa) and the ratio of internal and
815 external CO_2 concentration (C_i/C_a) calculated from flux tower observations at 16 sites. Each
816 point represents at least 10 days of data. Sites are coloured by IGBP vegetation type (blue=ENF,
817 brown=DBF/EBF green=CRO/GRA, black=WSA; see caption Table 1 for meaning of codes).

818

819