Supporting Information for

Cotton yield estimate using Sentinel-2 data and an ecosystem model over the Southern US

Submitted to *MDPI Remote Sensing*

Liming He¹* and Georgy Mostovoy¹

¹Laboratory of Environmental Model and Data Optima, Laurel, MD, USA, 20707

*Corresponding author: Liming He (liming.he@gmail.com; liming.he@emdos.com)
**S1 Summary of the BEPS model structure**

The Boreal Ecosystems Productivity Simulator (BEPS) - hourly version, is a process-based ecosystem model including water, energy and carbon budgets and soil thermal transfer modules (B. Z. Chen, J. M. Chen, & W. M. Ju, 2007; J. M. Chen, Liu, Cihlar, & Goulden, 1999; J. M. Chen et al., 2012; He, Chen, Liu, Bélair, & Luo, 2017; He et al., 2014). In this model, gross primary productivity (GPP) is modeled by scaling Farquhar's leaf-level biochemical model (Farquhar, Caemmerer, & Berry, 1980) up to the canopy level using a "two-leaf" approach (J. M. Chen et al., 1999; Norman, 1982). The bulk stomatal conductances of the sunlit and shaded leaves for water vapor and CO₂ are calculated using a modified Ball-Woodrow-Berry (BWB) stomatal model (Ball, Woodrow, & Beny, 1987). The Penman–Monteith equation (Monteith, 1965) is used to calculate the evaporation of intercepted water from the canopy and the ground surface, and canopy transpiration from sunlit and shaded leaves is computed following Y. P. Wang and Leuning (1998). The soil water dynamics is governed by the Richards equation (B. Chen, J. M. Chen, & W. Ju, 2007). The soil profile is stratified in five layers with depths of 0.05 m, 0.10 m, 0.20 m, 0.40 m, and 1.2 m from top layer to bottom layer. In BEPS, the influence of soil water on GPP is modeled through the modified BWB equation following G. B. Bonan (1995) and Weimin Ju et al. (2006).

Although BEPS was initially developed for boreal ecosystems, it has been expanded and used for temperate and tropical ecosystems in Asia (Matsushita & Tamura, 2002; Matsushita, Xu, Chen, Kameyama, & Tamura, 2004), China (Feng et al., 2007), Germany (Q. Wang et al., 2004), and other global applications (J. M. Chen et al., 2012; Z. Chen et al., 2017; He et al., 2018; He et al., 2017; Luo et al., 2018).

We summarize a few parts of BEPS that are related to the GPP modeling in detail below (He et al., 2014).
Supplementary Figure 1. (Fig. S1) A diagrammatic sketch for the BEPS model.

Background: the BEPS model

S2 Photosynthesis

The canopy-level photosynthesis ($A_{\text{canopy}}$) is simulated as the sum of the total photosynthesis of sunlit and shaded leaf groups (J. M. Chen et al., 1999):

$$A_{\text{canopy}} = A_{\text{sun}} (g_{sc\_sun}) L_{sun} + A_{\text{sh}} (g_{sc\_sh}) L_{sh}$$  \hspace{1cm}(1)

where the subscripts "sun" and "sh" denote the sunlit and shaded components of the photosynthesis ($A$) and leaf area index (LAI, or $L$). $g_{sc}$ is the stomatal resistance for carbon molecules. The sunlit and shaded LAI are separated by (J. M. Chen et al., 1999; Norman, 1982):

$$L_{sun} = 2 \cos \theta \left( 1 - e^{-0.5 \Omega / \cos \theta} \right)$$

$$L_{sh} = L - L_{sun}$$  \hspace{1cm}(2)

where $\theta$ is the solar zenith angle, $\Omega$ is the clumping index.

The net rate of CO$_2$ assimilation (either $A_{\text{sun}}$ or $A_{\text{sh}}$) is calculated as (Farquhar et al., 1980):

$$A = \min \left( A_{\text{sun}}, A_{\text{sh}} \right) - R_d$$  \hspace{1cm}(3)

$$A_{\text{c}} = V_{\text{max}} f_T \left( T_i \right) \frac{C_i - \Gamma}{C_i + K_c \left( 1 + O_i / K_o \right)}$$  \hspace{1cm}(4)
where $A$, $A_c$, and $A_j$ are the net photosynthetic, Rubisco-limited and light-limited gross photosynthetic rates $\mu$mol m$^{-2}$ s$^{-1}$, respectively. $R_d$ is the daytime leaf dark respiration, $V_{c max}$ is the maximum carboxylation rate at 25 °C ($V_{c max, sun}$ and $V_{c max, sh}$ for sunlit and shaded leaves, respectively). $J_{max}$ is the electron transport rate at 25 °C. $C_i$ and $O_i$ are the intercellular CO$_2$ and oxygen concentration, respectively. $\Gamma$ is the CO$_2$ compensation point without dark respiration, $K_c$ and $K_o$ are the Michaelis-Menten constants for CO$_2$ and oxygen respectively. $I$ is the incident photosynthetically active photon flux (mmols m$^{-2}$ s$^{-1}$). $f_V(T_l)$ and $f_J(T_l)$ are the leaf temperature ($T_l$) response functions for $V_{c max}$ and $J_{max}$ respectively. In the model, the $J_{max}$ is estimated from $V_{c max}$ (Medlyn et al., 1999):

$$J_{max} = 2.39 \cdot V_{c max} - 14.2$$

In the current BEPS, $f_V(T_l)$ and $f_J(T_l)$ share the same formula:

$$f_T = \frac{h_{kin} \cdot e^{ \frac{eakin(T_l - T_{opt})}{h_{kin}T_l}}} {h_{kin} - eakin \cdot \left(1 - e^{ \frac{h_{kin}(T_l - T_{opt})}{h_{kin}T_l}}\right)}$$

Where, $T_{opt}$ (301 K) is the optimum temperature for maximum carboxylation, and maximum electron transport, $r_{guc}$ (universal gas constant) = 8.314 J mole$^{-1}$ K$^{-1}$, $h_{kin}$ is the enthalpy term (200000.0 J mol$^{-1}$), $eakin$ represents the activation energy for electron transport, or carboxylation (55000.0 J mol$^{-1}$).

**S3 N-weighted $V_{c max}$ and $J_{max}$ for sunlit and shaded leaves**

The N-weighted $V_{c max}$ is derived according to J. M. Chen et al. (2012):

$$V_{c max, sun} = V_{c max,0} N_0 \frac{k \left[1 - e^{-\left(k_n + k\right)\frac{L}{d}}\right]}{\left(k_n + k\right) \left(1 - e^{-kL}\right)}$$

$$V_{c max, sh} = V_{c max,0} N_0 \frac{\frac{1}{k_n} \left(1 - e^{-k_nL}\right) - \left(1 - e^{-\left(k_n + k\right)L}\right)}{L - 2 \cos \theta \left(1 - e^{-kL}\right)}$$
where $V_{cmax,0}$ is the leaf maximum Rubisco capacity at the top of the canopy at 25°C, $\chi_a$ is the ratio of measured Rubisco capacity to leaf N (Dai, Dickinson, & Wang, 2004; dePury & Farquhar, 1997), $N_0$ is the N content at the top of the canopy; $k = G(\theta)\Omega / \cos \theta$, $G(\theta)$ is the projection coefficient, usually taken as 0.5 for spherical leaf angle distribution, $k_n$ is the leaf N content decay rate with increasing depth into the canopy, taken as equal to 0.3 after dePury and Farquhar (1997).

**S4 Surface evaporation and Canopy level transpiration**

The latent heat (LE) is simulated as:

$$LE = \lambda \left( T + E_i + E_g \right)$$  \hspace{1cm} (9)

where $\lambda$ is the latent heat of vaporization. $T$ is the transpiration rate from canopy (kg m$^{-2}$ s$^{-1}$), $E_i$ and $E_g$ are evaporation rates of intercepted water from canopy and ground surface (kg m$^{-2}$ s$^{-1}$), respectively.

The canopy level transpiration is obtained by:

$$T = T_{sun}(g_{r_{sun}})L_{sun} + T_{sh}(g_{r_{sh}})L_{sh}$$  \hspace{1cm} (10)

where $T_{sun}$ and $T_{sh}$ are the average transpiration rates for sunlit and shaded leaves, respectively.

The nonlinear relationship between $T_{sun}$ ($T_{sh}$) and $L_{sun}$ ($L_{sh}$) is considered in the parameters used to calculate $T$. $g_s$ is stomatal resistance for water molecules. $g_s/g_{sc} = 1.6$. Following Y. P. Wang and Leuning (1998), transpiration from sunlit leave is calculated as (W. Ju, Wang, Yu, Zhou, & Wang, 2010):

$$T_{sun} = \frac{D_a + \Delta(T_{r_{sun}} - T_s)}{r_{sun}} \frac{\rho C_p}{\gamma}$$  \hspace{1cm} (11)

where $D_a$ is the atmospheric vapor pressure deficit (kPa), $\Delta$ is the rate of change of the saturated vapor pressure with temperature (kPa °C$^{-1}$). $T_{r_{sun}}$ and $T_s$ are temperatures at sunlit leaf surface and air temperature (°C), respectively. $\rho$ is the air density (kg m$^{-3}$). $C_p$ is the specific heat of air at constant temperature (1010 Jkg$^{-1}$°C$^{-1}$), and

$$r_{sun} = r_a + \frac{1}{g_{r_{sun}}}$$  \hspace{1cm} (12)
where $r_a$ and $r_b$ are aerodynamic and boundary layer resistance ($s m^{-1}$), respectively, and $\gamma$ is the psychrometric constant (kPa °C$^{-1}$). To calculate $T_{sh}$, $T_{s,sh}$ (temperature at shaded leaf surface) and $g_{s,sh}$ are used to replace $T_{s,sun}$ and $g_{s,sun}$ in eq. (11) and (12).

The evaporation from soil $E_g$ is estimated using the Penman–Monteith equation (Monteith, 1965):

$$\lambda E_g = \frac{\Delta \left(R_g - 0\right) + \rho C_p VPD_g g / r_{a,g}}{\Delta + \gamma \left(1 + r_{soil} / r_{a,g}\right)}$$

(13)

where $R_g$ is the net radiation in the ground, $VPD_g$ is Vapor pressure deficit at the ground level, $r_{a,g}$ is the aerodynamic resistance of ground surface, $r_{soil}$ is the soil resistance for evaporation. In Sellers et al. (1996),

$$r_{soil} = \exp\left(8.2 - 4.2 \frac{\theta_1}{\theta_s}\right)$$

(14)

where $\theta_1$ is volumetric soil VWC in first layer ($m^3 m^{-3}$), and $\theta_s$ is value of $\theta$ at saturation ($m^3 m^{-3}$).

The $r_{soil}$ from Sellers et al. (1996) is a rough estimate that is derived from bare soil surface (Sellers, Heiser, & Hall, 1992). The evaporation can be overestimated if this equation is used since it does not consider the organic layer in the soil horizons. In BEPS, we used $4 * r_{soil}$ in the BEPS model.

The evaporation from intercepted water from sunlit and shaded leave $E_l$ are estimated similarly using eq. (13) to (14), but without the term for stomatal resistance (i.e., $r_s=0$).

**S5 Simulation of stomatal closure with rising CO₂ concentration in BEPS.**

Leaf stomata control the exchanges of water vapor and CO₂ between plants and the atmosphere. Under high atmospheric CO₂ concentration, stomatal density and hence conductance may decrease (Franks & Beerling, 2009). BEPS inherits the Ball-Woodrow-Berry (BWB) equation to model stomatal conductance ($g_s$, $\mu$mol m$^{-2}$ s$^{-1}$) (Ball et al., 1987):

$$g_s = g_0 + m \cdot h_s \cdot p \cdot \frac{A}{C_s}$$

(15)

where $g_0$ is a small value, the stomatal conductance at the light compensation point, $m$ is a plant species dependent coefficient, $h_s$ is the relative humidity at the leaf surface, $p$ is the atmospheric pressure, $A$ is the photosynthesis rate, and $C_s$ is the molar fraction of CO₂ at the leaf surface.
The important influences of soil water on \( g \) and \( A \) are not mechanistically included in the original BWB formulation. Following G. B. Bonan (1995) and Weimin Ju et al. (2006), we modify it as follows:

\[
g_x = g_0 + f_w \cdot m \cdot h_s \cdot p \cdot \frac{A}{C_s}
\]  

where \( f_w \) is a soil water stress factor, which we assume to be a function of soil water content. In Weimin Ju et al. (2006), the \( f_w \) is modeled as:

\[
f_w = \sum_{i=1}^{n} f_{wi} w_i
\]

where \( f_{wi} \) is the soil water availability factor in layer \( i \), and calculated as:

\[
f_{wi} = \frac{1.0}{f_i(\psi_i) f_i(T_{s,i})}
\]

where \( f_i(\psi_i) \) is a function of matrix suction \( \psi_i(m) \) (Zierl, 2001):

\[
f_i(\psi_i) = \begin{cases} 
1.0 + \left[ \frac{\psi_i - 10.0}{10.0} \right]^5 & \psi_i > 10 \\
1.0 & \text{else}
\end{cases}
\]

where \( \alpha \) is suggested to be a function of plant type (J. M. Chen et al., 2012).

The effect of soil temperature on soil water uptake is described as follows (Gordon B. Bonan, 1991):

\[
f_i(T_{s,i}) = \begin{cases} 
\frac{1.0}{1 - \exp\left( t_1 T_{s,i}^{t_2} \right)} & T_{s,i} > 0 \\
\infty & \text{else}
\end{cases}
\]

where \( t_1 \) and \( t_2 \) are two parameters determining the sensitivity of water uptake by roots to soil temperature. In the BEPS, \( t_1 = -0.02 \) and \( t_2 = 2.0 \).

To consider the variable soil water potential at different depths, \( w_i \) is calculated as:
where $R_i$ is the root fraction in layer $i$.

Apparently, $g_s$ will increase with $A$ (due to increase in photosynthetically active radiation (PAR) and/or $V_{c_{\text{max}}}$) assuming there is no change in $f_w$, $m$, $h$, $p$, and $C_s$.

The BWB equation can simulate the stomatal closure due to CO$_2$ fertilization. Assuming that there is no change in $f_w$, $m$, $h$, $p$, $V_{c_{\text{max}}}$ and PAR, there is an associated increase in intercellular CO$_2$ concentration ($C_i$) for an increase in $C_s$. Since $A$ is often limited either by Rubisco or by Electron-transport rate, the increase in $A$ will not be proportional to $C_i$, or in other words, the ratio of $A$ to $C_i$ will remain the same or decrease with rising $C_i$. As a result, the $g_s$ in the left side of BWB equation will remain the same or decrease (leading to stomatal closure) with rising $C_s$ (Baldocchi, 1994).

The BWB equation is used in many climate models, such as those in Coupled Model Intercomparison Project Phase 5 (CMIP5, http://www.nature.com/ngeo/journal/v6/n6/fig_tab/ngeo1801_T1.html) and TRENDY (Sitch et al., 2008) to study the global transpiration decrease (or increase of water use efficiency) due to CO$_2$ fertilization (Frank et al., 2015; Swann, Hoffman, Koven, & Randerson, 2016).

**S6 Calculations of radiation at Sunlit- and Shaded- leaf groups**

We refer to “Appendix A. Algorithms for net radiation of vegetation and ground surface” by B. Chen et al. (2016) for radiation calculation.

**S7 Calculations of Sunlit- and Shaded- leaf temperatures**

For a sunlit or shaded leaf, its temperature ($T_l$) is calculated as below during an iteration.

$$T_l = T_a + \frac{R_n - VPD_a \cdot \rho_a \cdot C_{pca} \cdot p^*}{\rho_a \cdot C_{pca} \cdot (G_h + \Delta \cdot p^*)}$$  \hfill (22)

where, $T_a$ is the air temperature in °C, $R_n$ is the net radiation of sunlit- or shaded- leaf calculated from S1.6, VPD$_a$ is water vapor deficit at the reference height, $\rho_a$ is the density of air at 0 °C, $C_{pca}$ is specific heat of moist air above the canopy, $G_h$ is the total conductance for heat transfer
from the leaf surface to the reference height above the canopy, $\Delta$ is the rate of change (slope) of the saturated vapor pressure with temperature (kPa °C$^{-1}$),

$$p^* = \frac{G_w + G_{ww} \cdot (X_{cs} + X_{cl})}{\text{psychrometer}}$$

where, $G_w$ is the total conductance for water from the intercellular space of the leaves to the reference height above the canopy, $G_{ww}$ is the total conductance for water from the surface of the leaves to the reference height above the canopy, Psychrometer is the psychrometric constant (0.066), $X_c$ and $X_s$ are the fractions of canopy covered by liquid water and snow.
S8 Forcing data and model parameters.

Climate reanalysis data are the outputs of an Earth system model that assimilates various archived observations. Global reanalysis data are the best available datasets for this study. MERRA-2 (Modern-Era Retrospective Analysis for research and Applications, Version 2) data from GSFC, NASA are used to drive BEPS to simulate GPP and ET in 2017 (Rienecker et al., 2011). The data have a spatial resolution of 0.625° (longitude) by 0.5° (latitude) and a temporal resolution of one hour. To drive BEPS, relative humidity, wind speed, and air temperature at 2 m above the surface, surface atmosphere pressure and incoming solar shortwave flux, and total precipitation at the surface level are spatially interpolated to the 20 m grid. The precipitation data from MERRA are corrected by global gauge-based NOAA Climate Prediction Center "Unified" (CPCU) precipitation product (CPCU). Recent validation suggests that MERRA2 datasets have relative small errors comparing to a few other reanalysis datasets (Draper, Reichle, & Koster, 2018; Eyre & Zeng, 2017; Reichle, Draper, et al., 2017; Reichle, Liu, et al., 2017; Simmons et al., 2017).

To simulate the CO2 fertilization effect, the CO2 concentration data are from https://www.esrl.noaa.gov/gmd/ccgg/trends/global.html.

S9 Previous validations of BEPS

Recent validations of GPP against eddy covariance measurements suggest that BEPS can explain more than 80% of the daily GPP variance at flux tower sites (Gonsamo et al., 2013; Sprintsin, Chen, Desai, & Gough, 2012). When soil water stress is properly addressed, BEPS explains 56-90% of the hourly GPP variance for maximum LAI values ranging from 2.1 to 8 (B. Chen et al., 2016). In 2018, the BEPS-simulated GPP is validated against eddy covariance measurements from 124 flux tower sites (FLUXNET2015 Dataset in Tier 1; http://fluxnet.fluxdata.org/) at the site level; validation suggests that BEPS simulates annual GPP well with a coefficient of determinations (R^2) of 0.81, a RMSE of 347 g C m^-2 yr^-1, and a bias of 172 g C m^-2 yr^-1 (He et al., 2018).


