Modeling profiles of micrometeorological variables in a tropical premontane rainforest using multi-layered CLM (CLM-ml)

Jaeyoung Song[1,2], Gretchen R. Miller[1]*, Anthony T. Cahill[1], Luiza Maria T. Aparecido[3,4], and Georgianne W. Moore[3]

[1]Department of Civil Engineering, Texas A&M University, 3136 TAMU, College Station, TX, 77843, United States
[2]Innovative Meteorological Research Department, National Institute of Meteorological Sciences, 33 Seohobuk-ro, Seogwipo-si, Jeju, 10551, Republic of Korea
[3]Department of Ecosystem Science and Management, Texas A&M University, 2138 TAMU, College Station, TX, 77843, United States
[4]School of Earth and Space Exploration, Arizona State University, Tempe, AZ, 85287, United States

Key Points:

- CLM-ml v0 was updated and compared to observations across a vertical profile in the tropical montane rainforest.
- Multi-layered CLM reasonably estimates key micrometeorological variables.
- Vertical leaf-distributions and turbulence play a significant role in land-atmosphere interactions.

Corresponding author: Gretchen R. Miller, gmiller@tamu.edu

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2020MS002259.

Abstract

This study updates the multi-layered Community Land Model (CLM-ml) for hillslopes and compares predictions from against observations collected in tropical montane rainforest, Costa Rica. Modifications are made in order to capture a wider array of vertical leaf area distributions, predict CO_2 profiles, account for soil respiration, and adjust wind forcings for difficult topographic settings. Test results indicate that the modified multi-layer CLM model can successfully replicate the shape of various micrometeorological profiles (humidity, CO_2, temperature, and wind speed) under the canopy. In the single-layer models (CLM4.5 and CLM5), excessive day-to-night differences in leaf temperature and leaf wetness were originally noted, but CLM-ml significantly improved these issues, decreasing the amplitudes of diurnal cycles by 67% and 47%. Sub-canopy considerations, such as canopy shapes and turbulent transfer parameters, also played a significant role in model performance. More importantly, unlike single layer models, the results that CLM-ml produces can be compared to variables measured within the canopy to provide far more detailed diagnostic information. Further observations and model developments, aimed at reflecting surface heterogeneity, will be necessary to adequately capture the complexity and the features of the tropical montane rainforest.

Plain Language Summary

This study is to improve and examine a multi-layered land-surface model for a better understanding of the surface process and advanced future climate prediction. This study was made through comparison with a single-layer model and with site observations about a tropical montane rainforest in Costa Rica. To apply the multi-layer model at this site, we updated a vertical leaf distribution and turbulence scheme and added the CO_2 profile and soil respiration scheme. The study showed the multi-layer model could more correctly reflect the site uniqueness (e.g., extremely wet) and the complexity by hillslope, compared to the single-layer model.

1 Introduction

Tropical forests play a key role in determining global and regional climate, and their associated land-surface processes are critical to the Earth system. Due to their significance for the global water cycles (K. Zhang et al., 2010; Choudhury & DiGirolamo, 1998) and climate cycles (Huntingford et al., 2013; Beer et al., 2010), improved modeling of tropical regions is vital for the accurate prediction of future climate and for the assessment of its impact on climate change. In the terrestrial biosphere, tropical forests house 25% of the carbon stocks and account for 33% of net primary production (NPP) (Bonan, 2008). While tropical forests comprise only 16% of the global surface, they produce 33% of terrestrial evapotranspiration (ET, 1,000-2,200 mm per year), 70% of which comes from transpiration (Schlesinger & Jasechko, 2014; Kume et al., 2011; Fisher et al., 2009; Loescher et al., 2005; Sheil, 2018). In humid tropical regions, hydrological processes are also markedly categorized by uniform warm temperatures, large inter-annual and spatial variability of moisture cycle, and high annual rainfall. Energy exchanges between the land and the atmosphere are enhanced by low albedo and high evaporative cooling (Wohl et al., 2012; Bonan, 2008). Anthropogenic changes (e.g., deforestation, climate change) can impact both the tropical forest itself and extratropical regions (D. Lawrence & Vandecar, 2014). Therefore, reliable prediction and precise evaluation of such effects must be addressed using Earth system models.

Unfortunately, land surface models (LSMs) do not yet adequately capture land-atmosphere interactions (Cai et al., 2014; Wang et al., 2014; D. M. Lawrence et al., 2011; Oleson et al., 2010). The predictions generated by LSMs are subject to significant errors, particularly in tropical regions (Bonan et al., 2011, 2012). The causes of this model error are not yet certain, making it an important area of ongoing study. D. M. Lawrence

et al. (2011) explored CLM4.0 (Oleson et al., 2010) results, comparing them to observations such as sensible and latent heat flux data from FLUXNET (Baldocchi et al., 2001). They recognized that the newer version (4.0) of CLM enhanced its predictive abilities compared to the previous version for a range of sites across the globe, but it still showed low correlations in tropical areas. Bonan et al. (2011) again updated CLM4.0 by changing the biological parameters and the structure of radiative transfer model for the canopy system, which resulted in major improvements (CLM4.5 (Oleson et al., 2013)) but some variables, such as trace gas fluxes, are still overestimated in equatorial regions. Other models, such as the Joint UK Land Environment Simulator (JULES) and the Australian Community Land Surface Model (CABLE), have the same issue (H. Zhang et al., 2013; Slevin et al., 2017). Many studies assert that latent heat fluxes are largely correlated (≈87%) with net radiation in tropical sites (Andrews, 2016; Fisher et al., 2009; Hasler & Avissar, 2007; Loescher et al., 2005). However, such a high correlation between ET and net radiation is insufficient information for accurate prediction because ET is known to be the dominant process in the tropical forest (Song et al., 2020). Rather, we need to investigate and improve the detailed mechanisms affecting water-related processes and variables such as evaporation of water directly from leaf surfaces (Loescher et al., 2005; Kume et al., 2011), aerodynamic conductance (Shuttleworth, 1988; Loescher et al., 2005), and vapor pressure deficit (VPD) (Fisher et al., 2009; Kume et al., 2011).

Therefore, the elaborate partitioning of both net radiation and water fluxes should be a major goal for accurate prediction at tropical sites, particularly where the place has an extreme environment like Costa Rica: frequent heavy rainfall, woody, and steep terrain (Song et al., 2020). LSMs contain complex, intertwined sub-models within the energy balance and water balance, which makes partitioning of their components challenging to fully understand. LSM studies require extensive field-based data sets for verification and parameterization of each sub-model. Hence, it has been an essential task to develop accurate models of individual processes through their observations (e.g., photosynthesis, soil, root, transpiration, canopy water, etc.) (D. M. Lawrence et al., 2011; Bonan et al., 2011, 2012, 2014, 2018; Burns et al., 2018; Swenson & Lawrence, 2014). For tropical sites, hydrological processes (e.g., canopy interception or soil infiltration) need to be especially accurate, because they primarily affect ET and the energy partitioning at the canopy. Other elements, such as thermal flux, radiative transfers, biogeochemistry, and vegetation activity and structure, are likewise important because they are mutually dependent on hydrological processes.

In our previous study, we found that a single layer model (e.g., CLM4.5, CLM5) was frequently insufficient to represent surface conditions at this tropical forest site (Song et al., 2020). In many studies, the near-surface layer, between the soil surface and the top of the vegetation, is described by one or two uniform control surfaces in a single layer, formulations known as "Big-Leaf" models (Dai et al., 2004; Oleson et al., 2013). Applying such models to the energy budget in a single surface layer provides computationally efficient and tolerably accurate results for many study sites (Ryder et al., 2016). However, they cannot fully represent the response of trace gas or energy fluxes from the surface due to large differences in generating mechanisms such as vegetation growth, leaf trait diversity, turbulent transfer, and energy exchange within the canopy. Latent and sensible heat fluxes are sensitive to model structures and process-based parameters, but the Big-leaf models are too simplified to show such effects (Jiménez et al., 2011). The structure of the land surface models can be more important than input data for evapotranspiration (Schlosser & Gao, 2010; Ryder et al., 2016). Some studies also reported that Big-Leaf schemes could not properly predict fluxes of sensible and latent heat due to the absence of the vertical structure of canopy (Jiménez et al., 2011; Ogée et al., 2003; Pitman et al., 2009; Bonan et al., 2014). The main cause of this shortfall was the failure to partition incoming solar radiation as a function of height adequately; it is one of the most vital inputs needed to accurately simulate transpiration (TR), carbon uptake, and energy balance through dense canopies.

Due to the potential of multi-layered schemes, development in this area is ongoing for several LSMs in addition to CLM. However, most of them have similar functional forms, so the response of CLM should be sufficiently representative of other current LSMs. In a typical model, a light profile scheme is applied to predict net radiation at each layer, and a wind profile model is then used to estimate the magnitude of flux transfer between layers. Finally, each flux of interest is estimated based on a scalar mass-conservation equation about the vertical exchange of each species (e.g., heat, vapor, CO_2). However, the models differ in their internal formulations, which themselves have varying sub-methods and levels of complexity [Table 1]. For wind profiles or turbulent transfer schemes, some models use a numerical scheme (1st-order or higher-order closure) but others use an extended version of the Monin–Obukhov Similarity Theory model (MOST) (Harman & Finnigan, 2007; Leuning, 2000) [Table 1]. MOST, which has a closed-form solution, has been widely used for many single-layer land surface models such as Community Land Models (CLM5) (D. M. Lawrence et al., 2018) and the Noah-MP land surface model with multi-parameterization options (Niu et al., 2011).

Table 1. List of light and wind profile formulations used in multi-layer land surface models

Model	**Full Name**	**Radiative Transfer Models (Light Profiles)**	**Turbulence Models (Wind profiles)**
ACASA	Advanced Canopy-Atmosphere-Soil Algorithm (Pyles et al., 2000)	Meyers and Paw U (1987) method which has an approach similar to Campbell and Norman (2012)	3rd-order turbulence scheme (Meyers & Paw U, 1987)
MLCan	Multi-layer Canopy-Root-Soil model (Drewry et al., 2010)	Simple mechanical and iterative procedures from Campbell and Norman (2012)	1st-order turbulence closure or similar scheme based on K-theory (Massman & Weil, 1999; Katul et al., 2004)
APES	Atmosphere-Plant Exchange Simulator (Launiainen et al., 2015)	Mathematical scattering model from Zhao and Qualls (2005)	
ORCHIDEE-CAN	Organizing Carbon and Hydrology In Dynamic Ecosystems – CANopy (Ryder et al., 2016)	Gu et al. (1999) method, an alternate version of the two-stream radiative transfer model (Sellers et al., 1992)	
CLM-ml	Multi-layer CLM (Bonan et al., 2018)	Simple mechanical and iterative procedures as Campbell and Norman (2012) or the modified version of two-stream approximation method introduced by Bonan et al. (2011)	The extended version of MOST (Harman & Finnigan, 2007; Leuning, 2000)
SHAW	Simultaneous Heat and Water model (Flerchinger, 2000)	Flerchinger et al. (2009) method which has an approach similar to Zhao and Qualls (2005)	
MuSiCA	Multi-layer Simulator of the Interactions between a Coniferous stand and the Atmosphere (Ogée et al., 2003)	Gu et al. (1999) method, and alternate version of the two-stream radiative transfer model	

While incorporating a multi-layered scheme into CLM, Bonan et al. (2018) improved its turbulence model from MOST. The updated turbulence scheme, called the Roughness Sub-Layer model (RSL), was successfully applied to the CLM-based multi-layer model (CLM-ml). Despite the reasonable predictions of mean gradient and turbulent fluxes, MOST tends to fail within the roughness sublayer above or near canopy height (Bonan et al., 2018; Harman & Finnigan, 2008, 2007). The RSL, on the other hand, can partially reflect canopy information such as Leaf Area Index (LAI) and approximate in-canopy mass flux rates. By comparing MOST with RSL using CLM-ml, Bonan et al. (2018) revealed that the turbulent transfer scheme is a key element to determine a model's performance. Their study highlighted that the update of in-canopy structure (single to multiple layers) in LSMs reduced known bias in sensible and latent heat flux, GPP (gross primary production), and turbulent transfer itself (Bonan et al., 2018).

Despite the updates to CLM and other LSMs, the multi-layer scheme has not been sufficiently verified against field data due to the lack of study sites having profile observations (Bonan et al., 2018). Moreover, observations in geographically complex areas are rare (Song et al., 2020), although montane tropical regions, like our study site, are also known to play an important role in regulating and collecting moisture in the atmosphere (Wohl et al., 2012). The study site presented here contains several extreme conditions (Song et al., 2020), in comparison to temperate sites, such as high humidity, high precipitation rates, and a steeply sloped surface. Furthermore, a wide range of micrometeorological observations - including vertical profiles of temperature, CO_2 concentrations, and water vapor deficit - are available at this site. This study will be a test case for tropical climates, but also it will provide useful insight into the importance of multi-layered schemes in land surface models of tall forests. It will also develop several model updates to more accurately capture land surface process.

In a departure from other studies, which have normally modeled a flat surface, we will examine the influences of a steep hillslope. This topographic relief mainly affects wind profiles and canopy configurations. The effect of the hillslope canopy cannot be simply up-scaled to the top horizontal boundary of the 3-D control volume, as is done for a flat, homogenous canopy. Moreover, to apply in-canopy complexity, two different methods are used to estimate the wind forcing value and the wind profile: (1) the RSL scheme which is already embedded in CLM-ml and (2) the numerical solution to first-order closure model, . Additionally, estimates of CO_2 emissions from soil and calculations to determine CO_2 concentrations in the canopy airspace are also added to CLM-ml. The use of multiple leaf area density (LAD) profiles and corresponding displacement heights (d) are investigated in order to capture the high canopy complexity of this site.

In our previous study of this site, CLM5 showed some improvements in daytime carbon and vapor fluxes and nighttime temperature and evaporation, as compared to CLM4.5. However, modeling efforts have not yet sufficiently resolved the overestimation issues which normally occur in the tropical forest. It also fails to represent the complex terrain, such as vertically overlapped canopies by steep hillslope (Song et al., 2020). In this study, we updated the CLM-ml model (e.g., wind profile functions, canopy shapes, and parameters) to better capture energy, trace gas fluxes, surface complexity, and vertical biometeorological profiles. Our objectives were three-fold:

1. To begin to represent surface complexity in the CLM-ml by applying more realistic vertical leaf distributions and adding a numerical wind-speed model;
2. To highlight the in-canopy variability of the forest and demonstrate the advantages of using a multi-layer scheme for complex land surfaces; and
3. To compare point-scale predictions of both single-layer CLM and CLM-ml against micrometeorological and flux measurements in a montane tropical rainforest in Costa Rica;

2 Methodology

2.1 Study Site and Micrometeorological Measurements

The study site is located at the Texas A&M University Soltis Center nearby San Isidro de Peñas Blancas in Costa Rica ($10°23'13''N, 84°37'33''W$, around 600 m above sea-level). The Soltis Center has a mean annual precipitation rate of 4200 mm, an average temperature of 24 °C, and an average relative humidity of 85% (Teale et al., 2014). This area is categorized as a premontane tropical rainforest. The canopy ranges in height between 24 and 45 m above steeply sloping terrain (Aparecido et al., 2016). Rainfall is common; over two-thirds of days have at least one rain event (Song et al., 2020).

The site has two different towers for biometeorological measurements, and their collective data were used for the simulation and comparison from mid-2014 to the end of 2016. The main weather tower (hereafter called "Met Tower") is situated over grass in a clearing at the edge of the forest. The walkup canopy access tower (hereafter called "Canopy Tower") is placed within the forest, on the eastern slope. The Canopy Tower measures a range of micrometeorological variables using a trace gas profile system (AP200, Campbell Scientific) with CO_2 and H_2O intakes and temperature sensors at eight discrete heights and an eddy-covariance (EC) system (LI-7200, LI-COR, Lincoln, NE; CSAT3, Campbell Scientific, Logan, UT) at 33 m. Other measured micrometeorological variables are tree transpiration rates (sap-flow), leaf wetness, net radiation, photosynthetically active radiation (PAR), and soil temperature, wind speed, carbon, and vapor fluxes. More detailed information about these measurements are provided by Song et al. (2020). For sap-flow sampling and installation, please refer to Aparecido et al. (2016).

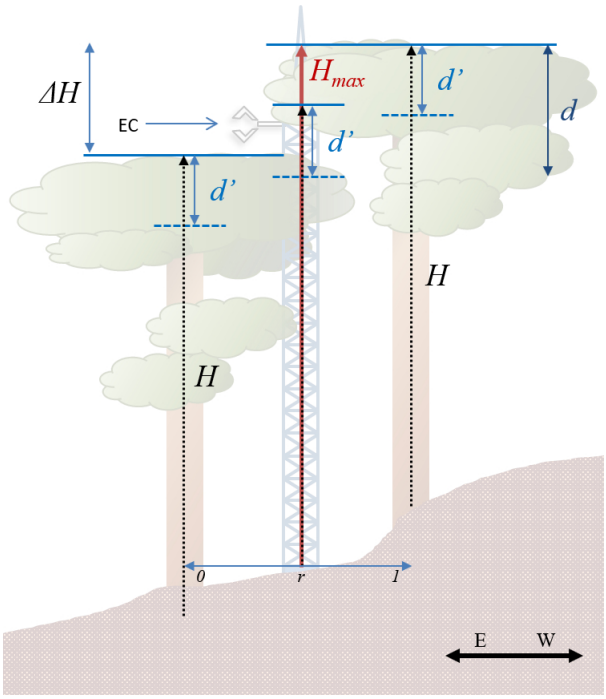

Figure 1. Conceptual schematic showing the canopy structure and instrumentation in relation to the surrounding hillslope. This figure is to help a reader imagine the dual-canopy concept between two adjacent trees at Canopy Tower. The symbol H denotes average canopy height across the forest stand, H_{max} is the maximum canopy height at the measurement location, r represents the Canopy tower location, 'EC' is the location of the sonic anemometer and infrared gas analyser, and d and d' are displacement heights from H_{max} and H. Predominant winds are from the north.

The Canopy Tower was positioned above the canopy but there is an emergent tree [Figure 1], leading to a large vertical opening between the two layers from approximately 30 to 40 m. Above the gap, the emergent tree upslope provides a substantial degree of shading; we see a 70% drop in photosynthetically active radiation (PAR) between the top of the downslope canopy (30 m) and above the emergent tree (44 m). This configuration also has some implications for the eddy-covariance measurements, which are less than ideal under these circumstances. The sonic anemometer and the infrared gas analyzer (IRGA) were located at 33 m, extending away from the tower and clear of obstruc-

tions in both the upwind and downslope directions [Figure 1]. Major winds occur parallel to the mountain, along the valley rather than over the slope, allowing us to capture fluxes, albeit under a narrowed set of ambient conditions. Thus, while these data are not necessarily sufficient for recording long-term, integrated measures of variables like gross primary production, they are appropriate for testing and validating models which can be assessed despite the presence of gaps caused by the interference of the emergent tree (Song et al., 2020). While using a footprint model would be optimal for screening the eddy-covariance measurements, none of the standard formulations are applicable as the surface is not homogeneous and is highly sloped. Instead, the data was deleted when the wind was not blowing in the correct direction based on the quality control flag (see section 2.4)

2.2 Site-Specific Meteorological Forcing Data

Our previous study (Song et al., 2020) compared CLM4.5/5 to our observations at the same site (the Soltis Center), and the same forcings were adopted here for CLM-ml (precipitation, wind speed, incoming solar radiation, temperature, air pressure, and relative humidity (RH)). These data were originally collected from the Canopy Tower in the forest, although precipitation data were only from the Met Tower. Missing data were gap-filled using Met Tower data, since meteorology data from the two towers are reasonably correlated with each other (e.g., R_0^2 =0.81 for wind speed, see Appendix A) and they provide similar results in the model. When both data sets had gaps, the missing data were replaced by available data from a randomly selected day from the same month and same time. Each 6 input forcing data has a missing rate from 5% to 23.5%. This gap-filling was to avoid simulation error, and these missing time steps were not used for any comparison or analysis.

Song et al. (2020) presumed that the forest has a flat surface and a uniform canopy height averaging around 36 m (Aparecido et al., 2016). This assumption made it possible to use the default canopy height ($\simeq$ 35 m) found in the satellite phenology (SP) mode in CLM. We need to note that the simulation results were not affected substantially when the forcing heights were varied (e.g., from 35 to 50 m), mainly because the vertical profile of wind speed with a short canopy reaches an inertial sublayer faster than with a tall tree. Within this layer, the vertical gradient of wind speed becomes very low, which results in a low mass transfer rate. This effect includes humidity and temperature values so they are not vertically sensitive within this height range.

Unlike previous studies, here we assumed that wind data was affected by both upslope and downslope canopies, which reach heights of 30.2 m (r = 0) and 43.7 m (r = 1) at the tower [Figure 1], respectively. The 13.5 m difference is due to the 45 degree slope angle in which the site sits; however, the approximate average height of the overstory canopy is around 39 m. The use of these values will be discussed in the next section. The change leads to higher maximum canopy height and results in a major difference in a simulation setup from the previous single-canopy approach. First, the wind speed at the sonic anemometer for EC system is now assumed to be within the canopy, no longer located above it. This means that the higher the altitude, the more dramatically the wind speed can increase. Therefore, it was necessary to extrapolate wind speed to a higher location (50 m in this study) based on the measured data (33 m).

From here, we developed 14 diverse sets of simulation cases to represent the interacting effects of canopy structure and wind [Table 2]. Of these, two mimic the previous study's settings (i.e., a single-layer with a 35 m the maximum canopy height and a 44 m forcing height in Song et al. (2020)). They may be directly compared to determine the influence of structural changes between CLM4.5/CLM5 and CLM-ml. The remainder of cases used the 50-m wind forcing height and two different wind profile models, but with a range of possible leaf area distributions (discussed below).

2.3 Model Development for Modeling Complex Forests

In this work, we present updates which added the consideration of the slope of the land surface and CO_2 profile simulation. These directly addressed three challenges presented by our study site, and others: (1) how to handle complicated canopy shapes in the model, (2) how to setup the wind forcing data and turbulence models, and (3) how to properly formulate boundary conditions for a CO_2 profiles scheme. As such, the influence of LAD distributions, maximum canopy height, and displacement heights on predictions of in-canopy micrometeorological variables was assessed. Since soil moisture was above the plant water stress point during this study and LAI was high, we assumed water supply to the vegetation was not limited and soil-related variables were not as influential as they typically are in drier, sunnier forests (Song et al., 2020). Thus, the most important factor in capturing the hillslope's effect is the ability to reflect the full vertical leaf distribution in places where trees significantly overlap. To achieve this, measured leaf area density (LAD) was simplified through a statistical distribution, and a turbulence scheme was updated to reflect LAD's shape. A detailed description of finding LAD from PAR measurements is presented in Song et al. (2020). The simulation setup and CLM-ml modifications needed to address these issues are explained below.

2.3.1 Canopy Shape

First, the canopy shape (i.e., the observed LAD) was determined using PAR sensor data and a light extinction model, the Beer-Lambert Law (Lalic et al., 2013; Maass et al., 1995). As expected, the data showed two major peaks in the LAD profile (Song et al., 2020), what we label here as a "dual-canopy" shape. However, CLM-ml does not include a dual-canopy distribution as an option, instead assuming a "single-canopy" shape defined by the equation:

$$f_{LAD,1}(z,H) = \frac{L}{H} \cdot f_{Beta}(z/H,p,q) + \frac{S}{H} \tag{1}$$

where $f_{LAD,1}$ $[m^2 \cdot m^{-3}]$ is the single-canopy model for leaf area density (LAD), z [m] is a height from the base of the tower, H [m] is canopy height, L $[m^2 \cdot m^{-2}]$ is leaf area index, S is stem area index, and p and q are shape parameters for the Beta distribution. Bonan et al. (2018) provide these parameters for grasses, crops, spruce trees, and pine trees, but we could find no sources specific to tropical forests. Pine trees ($p = 11.5$ and $q = 3.5$) were most similar to our observed LAD distribution: a tall tree with a dense canopy at the top.

We considered this to be potentially insufficient for modeling this forest, and thus developed a dual-canopy LAD model, which estimates the LAD at a given point (i.e., the tower) as a function of its proximity to nearby trees. It can be described using a mixed-distribution as:

$$\begin{aligned} f_{LAD,2}(z,H_d,r,\Delta H) = r \cdot f_{LAD,1}(z + (r-1)\cdot\Delta H, H_d) + \\ (1-r)\cdot f_{LAD,1}(z + r\cdot\Delta H, H_d) \end{aligned} \tag{2}$$

where $f_{LAD,2}$ $[m^2 \cdot m^{-3}]$ is the dual-canopy model for LAD, r [-] is a weighting parameter representing the distance of the tower from the downslope tree normalized by the distance between the two trees, H_d [m] is the height of dominant tree which is the same as H, the maximum canopy height H_{max} [m] between the two is estimated through $H_{max} = H_d + (1-r)\cdot\Delta H$, and ΔH [m] is the vertical distance between the top of the canopy on the downslope tree and the top of the canopy on the upslope tree [Figure 1]. Additional combinations could better mimic mid-story and sub-story structure in the canopy, however, here we assume the dual-canopy model is sufficient to represent the true canopy shape for this study. More detailed information on the formulation and parameterization of the LAD model may be found in Appendix B4.

Consequently, the dual-canopy LAD model results in a greater maximum height than the single-canopy model. In CLM4.5/5 and our previous study, a default canopy height of 35 m was obtained from global surface data, assuming flat terrain and a single-canopy shape. This down-slope canopy height, which was set right below the EC system, was also consistent with our expectations by comparison with the tower. However, when we instead consider a dual-canopy profile using a mixed Beta distribution, the fitted value for average canopy height was approximately H_d=39 m and maximum canopy height was H_{max}=43.7 m (with $r = 0.65$, and $\Delta H = 13.5$), due to the topographical slope effect . We assumed this to be most representative of reality.

2.3.2 Wind and Turbulence Scheme

We applied wind speed data from the canopy gap (EC 33 m) on the Canopy Tower as the forcing rather than use the usual forcing data measured at a standard height over a surface, like the Met Tower. The wind speeds of the two places were also sufficiently correlated [Figure A1], so that the choice did not significantly affect model results. On the other hand, the vertical profile cannot be the same, because the turbulent process over a grassland are likely much different than those over a sloped canopy surface. Given this issue, one promising idea was to use 33-m three-dimensional sonic anemometer (CSAT3) data. Then, this wind speed within the gap was used to track the full wind profile, reflecting the influence of the emergent tree via the vertical leaf distribution. In other words, the modified CLM-ml can estimate windspeed upward and downward from the middle height of the two canopies, not for downward tracking only.

Additionally, two different turbulence models were used in this study. While the previous wind profile model embedded in CLM-ml, RSL, is generally a practical and reliable update to MOST, it does not adequately reflect LAD distributions below the maximum canopy height H_{max} (See Appendix B2). Therefore, we added a first order closure model (FOC) to investigate the effects of different LAD distributions on performance. In our new mixed-model, the RSL scheme was applied above the maximum canopy height and the FOC equation was numerically solved in lower portions. These changes allowed us the advantages of both formulations. Despite the addition of an iterative solver, we were able to minimize the computational time in this test by applying wind profile u(z) from the previous time-step as the initial guess in subsequent time-steps.

The displacement height, where logarithmic wind profile goes to zero, was also updated to apply the change in canopy height. Here, we assumed that the calculation uses the spatially averaged height (H). In RSL, the displacement height was originally derived based on the drag at the centroid of the canopy (Harman & Finnigan, 2007). Since the slope effect made the drag area rhombus shaped, using the spatial average height (H) is more reasonable [Figure 1]. For instance, what we dub the "dual-canopy" concept takes into account a maximum height of 44 m from the tower base on one side and a 30 m height on the other. Normally, we have to use 44 m for 'd' value estimation. However, 44 m occurs because of slope effects, so it cannot represent the average canopy height. The estimation of displacement height is directly related to LAI and H which refers to the maximum canopy height H_{max} on the flat forest as:

$$H - d' = \frac{\beta}{C_d \cdot a} \tag{3}$$

where d' is displacement height using spatial average height (H), β consists of friction velocity and wind velocity as $\beta = u^* \cdot u(H)$, a is LAD ($= LAI \cdot H^{-1}$), and C_d is drag coefficient. The d' now refers to the distance from H, so final displacement height (d) becomes d' + H_{max} - H [Figure 1].

In addition to introducing the new LAD distributions, we modified the CLM-ml codebase in three significant ways by: 1) reformulating the wind speed profile scheme; 2) adding calculations to determine the in-canopy CO_2 profile; and 3) adding a repre-

sentation of soil fluxes to account for respiration. This soil respiration is particularly necessary for the CO_2 profile, as it acts as its lower boundary condition. The equations are discussed briefly here; more detailed information may be found in Appendix B.

The RSL model embedded in CLM-ml has an analytical solution in which the key function depends on the height ($u(z)$=$f(z)$), making its computation much simpler compared to the raw first-order closure model scheme, which requires a numerical solution method [See Appendix B2].

In this study, a first-order closure model was added to determine the wind profile and the eddy-diffusivity. This addition was needed to predict the effects of different LAD distributions because RSL (a closed-form approximation) cannot fully represent them. The model follows previous work (Launiainen et al., 2011; Katul et al., 2004; Drewry et al., 2010) and is given as:

$$K_m \frac{\partial^2 u}{\partial z^2} + \frac{\partial K_m}{\partial z}\frac{\partial u}{\partial z} - C_d a(z) u^2 = 0 \tag{4}$$

where K_m ($m^2 \cdot s^{-1}$) is the eddy diffusivity for momentum, u ($m^2 \cdot s^{-1}$) is wind speed, z is height, C_d is drag coefficient (0.25), and $a(z)$ ($m^2 \cdot m^{-3}$) is LAD as a function of height above the ground (Launiainen et al., 2011), see Appendix B1 for more detailed description. For $a(z)$, $f_{LAD,2}$ from Eq. (2) was used in this study. To estimate K_m, displacement height d (m) is necessary. The displacement height is usually set to $0.667 \cdot H_{max}$ but it varies in CLM-ml (Bonan et al., 2018). For this study, the displacement height is estimated either based on the average canopy height H when using a single canopy model or on the height of the nearest dominant tree H_d=H when using the dual canopy model. In the later case, H_d can be estimated from dual-canopy distribution function ($f_{LAD,2}$) fitted to the observations.

2.3.3 CO_2 Concentrations Profiles and Soil Respiration

The method to determine CO_2 concentrations as a function of height (i.e., CO_2 profile) follows a formulation similar to that used to compute water vapor in the Bonan et al. (2018) version. For CO_2, this equation was modified to:

$$\rho_m \frac{\partial C}{\partial t} - \frac{\partial}{\partial z}\left(\rho_m K_c(z) \frac{\partial C}{\partial z}\right) = \left[f_{c,sun}(z) f_{sun}(z) + f_{c,sha}(z) f_{sha}(z)\right] a(z) \tag{5}$$

where ρ_m is molar density ($mol \cdot m^{-3}$), C is CO_2 concentration ($\mu mol \cdot mol^{-1}$), t is time (s), K_c is scalar diffusivity, f_{sun} ($\mu mol \cdot m^{-2} s^{-1}$) is the fraction of sunlit leaves, f_c is photosynthesis flux, and the sum of terms found within the square brackets are sources and sinks representing leaf scale carbon assimilation. The value of K_c was assumed to be the same as K_m in this study, as previous literature reports that the ratio appears close to 1 (Launiainen et al., 2011).

Soil respiration R ($\mu mol \cdot m^{-2} s^{-1}$) was also added to CLM-ml as a source term at $z = 0$. As in Launiainen et al. (2011), the following expression was used:

$$R = R_{10} Q_{10}^{(T_g - 10)/10} \tag{6}$$

where R_{10} ($\mu mol \cdot m^{-2} s^{-1}$) and Q_{10} (-) are parameters which have values equal to 0.3 and 2 at 10 °C (Launiainen et al., 2011), and T_g is ground temperature (°C). These R_{10} and Q_{10} are also fitted using measured data in this site which have 2.4 and 1.7. This fitted soil respiration model (called as 'Q1' in this study) has high rates and sensitive to the ground temperature ($R \approx 4.5$ $\mu mol \cdot m^{-2} s^{-1}$ at 22 °C) compared to Launiainen et al. (2011) ($R \approx 0.7$ at 22 °C). This error for respiration is likely due to the relatively high ground temperatures found in all three models (CLM4.5/CLM5/CLM-ml), which leads

to excessive soil respiration rates. Therefore, the ground temperature for the 'Q1' simulation is directly predicted from the forcing temperature using the linear regression model, $R^2 = 0.75$.

2.4 Testing the Influence of the Sloped Surface on Flux Measurements

The eddy-covariance (EC) system was examined by identifying its possible representation height through a multi-layer model. Our previous study identified possible interference from the upslope emergent tree Song et al. (2020), although the EC system is located out of the Canopy Tower and above the lower (downslope) canopy. Such interference can have an impact on turbulence and makes it difficult to identify what the flux measurements actually represent. Moreover, carbon and vapor fluxes can be a relatively large scale compared to the wind/turbulence scheme, but the influence of upslope trees on their fluxes is also not fully understood. Here, the wind and its turbulence can be regarded as the forcing to transport species (fitting an Eulerian approach). The carbon and water vapor are the actual species which can move long distance (close to Lagrangian approach). The scale gets larger if the tree gets taller (more air space) as with our study site. This is mainly because the source for the fluxes depends on wind direction and the measurement height (Burba, 2013). Moreover, tracking the carbon or vapor can be complicated in the steep area. For instance, their source would mostly come from a downslope rather than an upslope if the kinetic energy is normally toward the atmosphere. Fortunately, at this site the predominant wind direction runs perpendicular (North), rather than parallel to the hillslope [Figure 1], which makes it possible to assume that the slope effect on the fluxes would be minimal. However, this approach is subject to a considerable amount of uncertainty.

Here, measured data were compared to simulation results to assess three possible hypotheses with the aid of the multi-layer model. We used high-quality data, in which a quality control flag was zero, based on post-processing performed in EddyPro version 6.2.0 (LI-COR, Lincoln, Nebraska, USA). First (H0), EC method can represent the full flux of the area, as was assumed in our previous study. Data comparison would be EC data versus total fluxes from CLM. Second (H1), EC method measures the partial flux of the area which does not contain the residual flux from the emergent tree above the EC system at 33 m: EC data versus fluxes at 33 m from CLM. This means the EC method cannot cover enough due to topographical complexity. Last (H2), EC data represents a mixture between the top and EC measurement height at the Canopy Tower. This test is possible due to a multi-layered model. If EC flux data was not falling between Top (H0) and 33 m flux (H1), then we can conclude that the model significantly over- or underestimates the fluxes. If H2 is accepted and CLM is overestimated, then we can conclude that the photosynthesis parameter causes the error, as observed in our previous study. If CLM is underestimated, it instead casts doubt on the predictions of low incoming solar radiation at lower canopy (light profile error), which would indicate that the light-extinction model may be too simplified for sloped canopies and possibly for sparse canopies having various heights.

2.5 Simulation Setup and Analysis

CLM-ml uses the CLM4.5 parameterization scheme, although some CLM5 parameters for interception and u* are available. The calculation of stomatal conductance uses water-use efficiency optimization and plant hydraulics introduced by (Bonan et al., 2014). The consideration of hydrological process was not specifically modified for sloped surfaces, but new the LAD formulation would affect how interception rates change across the hillslope. Also, the northern wind was dominant so the upslope and the downslope wind could be ignored.

CLM-ml runs with other common CLM modules (e.g., soil temperature) and default variables (Bonan et al., 2018). The original code (CLM-ml v0) (Bonan et al., 2018) was adapted to meet the requirements of this study. Then, the CLM-ml module, which only handles computations for canopy process, was inserted within CLM5 package as one of the various modules. The main loop in CLM5 ("clm_driver.F90") calls the CLM-ml module. For practical simulations, this study used a 30-node vertical grid, which provides sufficient resolution while remaining computationally efficient. Some formulations of CLM-ml, vertical leaf distribution, turbulence scheme, and CO_2 concentration scheme, were updated to reflect our complex site features and to ease inter-comparisons between single canopy models (CLM4.5/5), multi-layered CLM (CLM-ml), and our observations. This study added a CO_2 profile scheme since CLM-ml did not calculate it yet. A new leaf area distribution (LAD) and additional turbulence scheme was applied, to address complex terrain. Parameters and simulation setup for CLM-ml were default mode (the satellite phenology (SP) mode) and any extensions such as a demography model or biogeochemistry model were not active. Also, the default parameters, including total LAI and soil properties, were not modified if there is no significant improvement like the previous study (Song et al., 2020). This study assumes the two big trees can represent one grid cell. The two-step shape of leaf distribution is to simulate overlapping trees. This assumption makes sense because the two big canopies cannot meet at the same height on the steep hillslope. Considering more trees through the demography model would be a future study.

All simulation results and observations were compared using vertical profiles of relevant variables, grouped by time of day (daytime versus nighttime). The model was run at half-hour time steps, and the observations were averaged to match these outputs. All results shown here represent an average over the entire time range, from mid-2014 to the end of 2016. The up-scaled variables and fluxes were also compared with data grouped by wetness conditions (wet days versus dry days) in supplementary data sheets (each sheet's name is the same as each variable's name). The supplementary data sheets also provide the value of the mean and standard deviation for result figures, and the supplementary figures show their confidence interval line. Upscaling (or spatially normalizing in a vertical way) from the bottom to the top was also conducted for the canopy water and temperature to compare with the single-layered model (CLM). For example, upscaled-values (X) for diurnal variations about the two variables were estimated as:

$$X = \left(\sum_z LSAI_z \cdot x_z \right) / \sum_z LSAI_z \tag{7}$$

where z is heights (m) which represents each node, and $LSAI_z$ is the sum of leaf and stem area index (LAI_z + SAI_z, $m^2 \cdot m^{-2}$) at each node.

The simulation settings are listed [Table 2], where z_{max} is forcing height except wind, z_u is wind forcing height, H_{max} is maximum canopy height, d is displacement height, $H(d)$ canopy height for displacement calculation d, p and q are parameters for the Beta function, and r and ΔH is LAD distribution parameters. These parameters give different canopy shapes in simulations, which affects both fluxes and other meteorological variables. We use 'numerical' to denote when the turbulence model was changed from RSL to the first-order closure model; in the simulation name, 'N--' indicates the use of first order closure (FOC) model by numerical method and 'R--' refers to RSL model [Table 2]. Again, that replacement from RSL to FOC was made only for under the H_{max}, to preserve the stability concept. The proposed turbulence model for this study is a mixed model (RSL+FOC). The different canopy shapes are plotted in Figure 2, a single-canopy shape used the parameter for Pine trees ($p = 11.5$ and $q = 3.5$) and dual-canopy shape used fitted parameters ($p = 69.9$, $q = 8.7$, $r = 0.65$, and $\Delta H = 13.5$). These two distributions are our target shapes expecting that would provide better results as a sim-

plified distribution (-1C) and more complex distribution (- 2C), compared to uniform distribution (- FC).

In these case studies [Table 2], the wind profile was estimated based on the sonic anemometer data (u_{EC} at z_u=33 m) located in the upper two-thirds of the profile system [Figure 1; Figure A1]. Only in simulations #1-4, we assume that u_{EC} was z_u = 44 m in order to provide a more direct comparison with the assumptions in the standard CLM version. The first-order closure model Eq. (4), an ordinary differential equation (ODE), must be solved based on wind speed values at two boundaries - the sonic anemometer height (u_{EC}) and the ground (u_g=0). Then, the wind speed was iteratively estimated for successively higher points ($z > z_u$) until the maximum canopy height ($z <= H_{max}$) was reached.

Table 2. List of simulations and their naming conventions. In these names, 'R' and 'N' represent the RSL model and the numerical FOC model, respectively; '1C', '2C', and 'FC' indicate single, dual, and uniform canopy shape; 'o' refers to original settings, which means CLM-ml runs based on the same environment of CLM5; 'H' refers to $H = H_{max}$. Simulations without the 'H' sign are our major target settings since d is estimated by a new method. H_{max} is 43.7 m and the average H is 39 m which is used for new d. ΔH, p, q, and r are LAD shape parameters for Eq. (3) and Eq. (1)

#	Name	z_{max} / z_u	H_{max} / $H(d)$	Wind model	LAD	p	q	r	ΔH	R_{10}	Q_{10}
1	CLM5.0	44/44	35/35	MOST	-	-	-	-	-	-	-
2	CLM4.5	44/44	35/35	MOST	-	-	-	-	-	-	-
3	R1Co H35m	44/44	35/35	RSL	Single	11.5	3.5	-	-	0.3	2.0
4	R1Co H39m	44/44	39/39	RSL	Single	11.5	3.5	-	-	0.3	2.0
5	N1C H35m	50/33	35/35	Numerical	Single	11.5	3.5	-	-	0.3	2.0
6	N1C H39m	50/33	39/39	Numerical	Single	11.5	3.5	-	-	0.3	2.0
7	R1C H44m	50/33	43.7/43.7	RSL	Single	11.5	3.5	-	-	0.3	2.0
8	N1C H44m	50/33	43.7/43.7	Numerical	Single	11.5	3.5	-	-	0.3	2.0
9	R1C	50/33	43.7/39	RSL	Single	11.5	3.5	-	-	0.3	2.0
10	R2C	50/33	43.7/39	RSL	Double	69.9	8.7	0.65	13.5	0.3	2.0
11	N1C	50/33	43.7/39	Numerical	Single	11.5	3.5	-	-	0.3	2.0
12	N2C	50/33	43.7/39	Numerical	Double	69.9	8.7	0.65	13.5	0.3	2.0
13	NFC	50/33	43.7/39	Numerical	Flat	1	1	-	-	0.3	2.0
14	N1C Q1	50/33	43.7/39	Numerical	Single	11.5	3.5	-	-	2.4	1.7

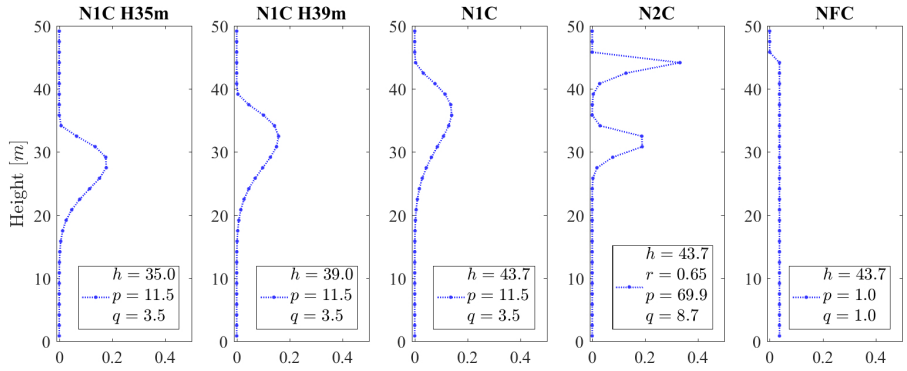

Figure 2. The five discrete LAD distributions from 'Numerical' run cases for FOC, as defined by the parameters p, q, r, and h=H_{max} [Table 2]. These LAD distributions are identical to RSL type runs. The sum of all vertical points is 1 in this distribution. The domain space was discretized into 30 computational nodes, shown here as dots.

3 Simulation Results and Discussion

Simulations were conducted based on parameters and settings listed in Table 2, for cross comparison between CLM, CLM-ml, and updated CLM-ml which includes a first-order closure model, LAD distributions, CO_2 concentration calculations, and a soil respiration scheme.

3.1 Energy Exchange

As expected, total net radiation (R_{net}) predictions were not significantly different between models (CLM4.5, CLM5, and CLM-ml) [Figure 3] because they were estimated from the same radiative transfer parameters, particularly albedo. The small difference in daytime values ($\leq 3\ W/m^2$) may have been due to the multi-layered radiative transfer scheme, which generates different degrees of energy exchanges at each layer, according to different canopy shapes [Figure 4a; Figure 4b; Figure 4c]. The impact of in-canopy LAD variation was small on net radiation but it was detected via vapor and carbon flux comparisons later. The nighttime net radiation decreased in CLM-ml, especially for cases with higher maximum canopy heights (see supplementary data spreadsheet). In this case, the wind turbulence model highly affected the energy exchange, similar to Bonan et al. (2018). In the daytime, as expected, both the net radiation profile and the PAR profile were naturally influenced by the light–extinction model but these results showed that they were also greatly affected by the canopy shape [Figure 4a; Figure 4b]. Consequently, this vertical variation of available energy affected physiological activities such as GPP and transpiration (TR), which had similar profile shape (see later section 3.5).

Simulated net radiation, PAR, and sensible heat profiles produced by the dual-canopy simulation (N2C) had two inflection points which occurred near the peak LAD values [Figure 4a; Figure 4b; Figure 4c]. We note that contrary to expectations, the amount of received energy was significantly different between the upslope tree and the downslope tree [Figure 4a] , although the magnitude of the two peaks are similar in the two-step function (N2C) [Figure 2]. This pattern arose because most energy can be absorbed by the top canopy. We need to note that the two canopies spatially overlap slightly, and the tower was placed at the overlapping place [Figure 1]. The energy received at this research site may have spatio-temporal variability depending on the angle of the sun and measurement location. Hence, this LAD profile contains measurement error because the

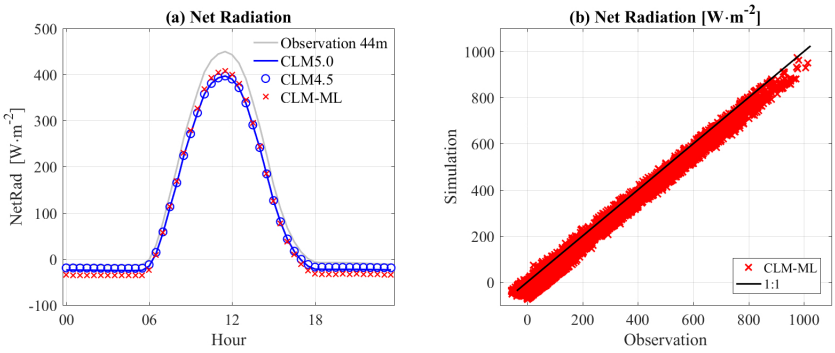

Figure 3. (a) Diel trends and (b) correlation between net radiation measured at 44 m on the Canopy Tower and predictions from CLM. The R^2 of CLM5.0, CLM4.5, and CLM-ML against the observations were 0.9653, 0.9664, and 0.9708.

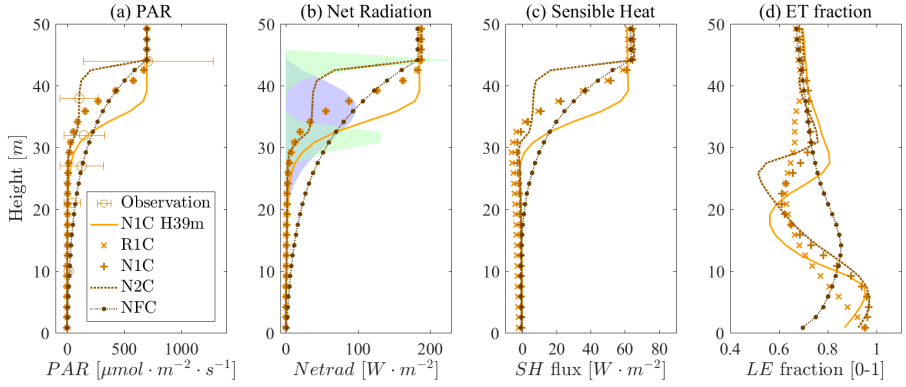

Figure 4. Profiles of (a) PAR, (b) net radiation, (c) sensible heat flux, and (d) daytime evaporative fraction as a function of height above the ground at the Canopy Tower. These profiles were simulated using the multi-layered version of CLM 4.5 (CLM-ml) using different leaf distributions and turbulence schemes [Table 2]. The purple and green area in (b) show the LAD distribution of N1C and N2C in Figure 2.

LAD profile was measured at a single location and cannot fully capture spatial variability.

Despite the inconsistency, the observations of the PAR profile matched the values modeled using the dual-canopy shape (N2C) better than using the other shapes [Figure 4a]. Here, PAR profile was highly correlated with the net radiation profile [Figure 4b], which is a critical variable in any LSM. Hence, the precise vertical layering of vegetation (LAD) is essential for the model accuracy.

Additionally, results showed even a very thin top layer of leaves can contribute disproportionally to the energy exchange despite being a small fraction of total LAI. The upper layers from 42 m to 44m accounts for most of the energy exchange, as shown in the net radiation profile [Figure 4b]. The ET profile has a very similar shape to PAR, net radiation, and heat flux, so ET fraction was used to show the portion of ET within net radiation. At this top layer, the ET fraction profile indicates that the ET was a main contributor to the energy exchange in this site. The ratio between evapotranspiration (ET) and sensible heat (SH), $ET{\cdot}(ET{+}SH)^{-1}$, was around 0.7 (Figure 4d) at at the topmost level of the canopy.

3.2 Wind Profile

The shape of the wind profile generated by the multi-layered scheme reasonably followed the observations within the canopy [Figure 5]. Applying the more realistic canopy shape and height (N2C, dual-canopy formulation) gave a more plausible wind profile, which had two transition points influenced by the two major canopies [Figure 5c]. Here, the shape of the canopy profile did not seem to yield large differences in wind speeds (i.e., N1C vs N2C) [Figure 5c]. However, their impacts would be shown later in other variables such as carbon and vapor fluxes. This analysis also indicated that the change of the canopy heights produced a notable impact on the wind profile [Figure 5a]. The different heights for these test cases were the mean canopy height of 35 m from the global data set normally used by CLM5, the mean canopy height of 39 m from Eq. (2), and the actual maximum height of 44 m due to the slope effect. The change of displacement height (d) also resulted in different profiles especially with RSL model, R1C H44m vs. R1C [Figure 5b]. We note that RSL was sensitive to parameters related to the displacement height rather than the canopy shape. On the other hand, the first-order closure (FOC) model, 'Numerical' scheme, was more sensitive to canopy shape [Figure 5c]. The two methods, FOC (N1C) and RSL (R1C), tended to generate slightly different wind profiles. Also, this small difference had little impact on the final fluxes simulated at the top, which will be discussed in later comparisons. However, the important point is that the influence of the canopy shape change can be seen only through the numerical FOC method.

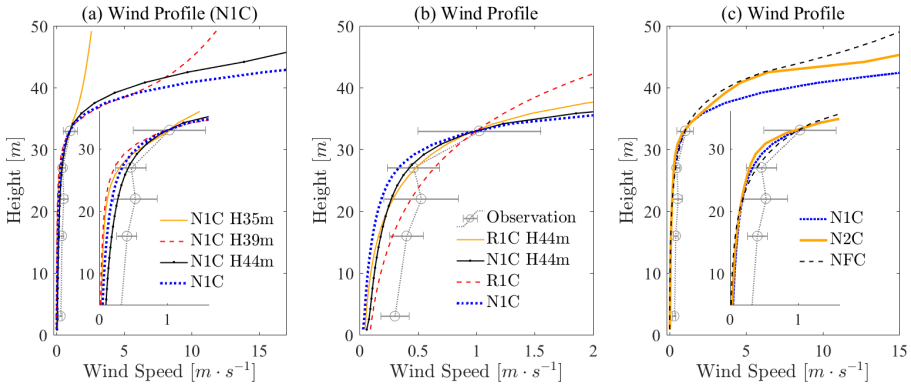

Figure 5. Variations in wind speed prediction as a function of height above the ground at the canopy tower according to (a) canopy height, where H indicates H_{max}=$H(d)$ [Table 2]; (b) Wind turbulence model and method of displacement height calculation; and (c) LAD distribution. Circles indicate observations, and prediction errors were estimated and shown in supplementary data.

3.3 Leaf Wetness

The multi-layered model significantly improved the prediction of leaf wetness [Figure 6]. The increase in nighttime evaporation had an important role in moderating the over-fluctuation in diel variation compared to the previous study (Song et al., 2020). Also, the higher canopy height in CLM-ml tended to follow measured values well. In overall results, the lower portions of the profile, which also had low leaf areas, tended to hold more water compared to observations. This was possibly caused by low net radiation at the lower heights as shown in the profile shape of net radiation. We could not visually identify the significant impact between two different turbulent transfer schemes (e.g., R2C vs N2C) but the small difference could be detected in the nighttime (8.9% vs 9.9%) and wet day (15.1% vs 16.4%). Nonetheless, the multi-layer model provided more reasonable

results as displayed in the comparison using up-scaled leaf wetness [Figure 6]. Considerable measurement error might be expected from these sensors due to their shape and low spatial coverage, but they did provide insight into diurnal variations in wetness.

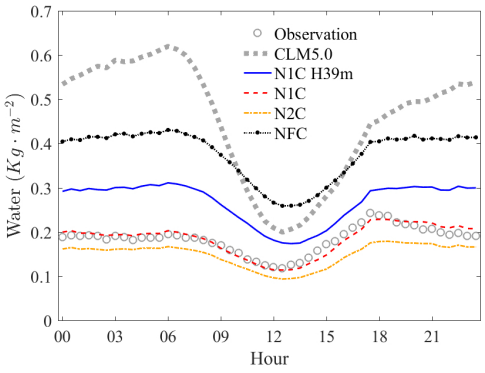

Figure 6. Diurnal variation in leaf wetness, up-scaled from profile results using Eq. (7). Here the power term (2/3) in CLM was not used for leaf wetness (y-axis). This leaf wetness represents the ratio between the current canopy water and the maximum canopy water $W_{max} = 0.1$ kg·m^{-2}. The LSAI is approximately ≈ 6 m^2·m^{-2}. The R^2 values for a 1:1 regression line of CLM5.0, N1C H39m, N1C, N2C, and NFC against the observation were -1.1275, 0.3019, 0.5544, 0.5561, and -0.2506, respectively.

3.4 CO_2 Concentration, Relative Humidity, and Air and Leaf Temperatures

Modeled profiles of air temperature, relative humidity, and CO_2 concentration showed general agreement with the observations [Figure 7; Figure 8; Figure 9]. The night-time profiles of air temperature and relative humidity were noticeably improved compared to the single-layer model. Most simulated profiles fell within one standard deviation of observed profiles, although some values at low canopy heights had larger errors. The R1Co case, in which the forcing wind speed began at 44 m similarly to single-layered CLM, showed lower air temperatures at the canopy like Bonan et al. (2018), which was opposite the behavior observed in the field.

The standard deviation of relative humidity and air temperature in the daytime was bigger than at nighttime. The site is very frequently rainy and wet. In the daytime, incoming radiation interacts with these the surface conditions to generate a frequent cooling effect and high relative humidity levels.

The results also indicated that these variables were mainly affected by turbulence-transfer parameters, especially by the maximum canopy height H_{max} and LAD distributions [Figure 7; Figure 8; Figure 9]. In particular, CO_2 concentration resulted in notably different outcomes by each separate case [Figure 7]. For instance, low canopy heights such as N1C H39m tended to have more errors especially in CO_2 profile. The flat LAD (NFC), which is not ball-shape distributions like N1C, resulted in different vertical profiles in daytime CO_2 concentration and air temperature, compared to other results.

Both the CO_2 profile and CO_2 flux were highly influenced by the soil respiration [Figure 12; Figure 7]. Applying parameters from Launiainen et al. (2011) resulted in reasonable predictions but applying our measured soil flux (N1C Q1) made the results substantially deviate from our observations. However, we could not conclude yet which is

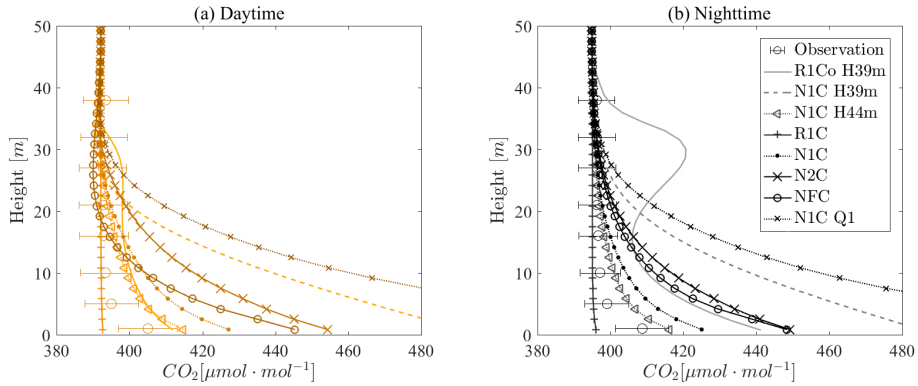

Figure 7. CO_2 concentration profiles measured by the AP200 system (open circles with error bars) and modeled (lines) during the (a) daytime and (b) nighttime.

correct, due to the significant error of GPP prediction in CLM-ml that still exists and the high spatio-temporal variability of of soil fluxes.

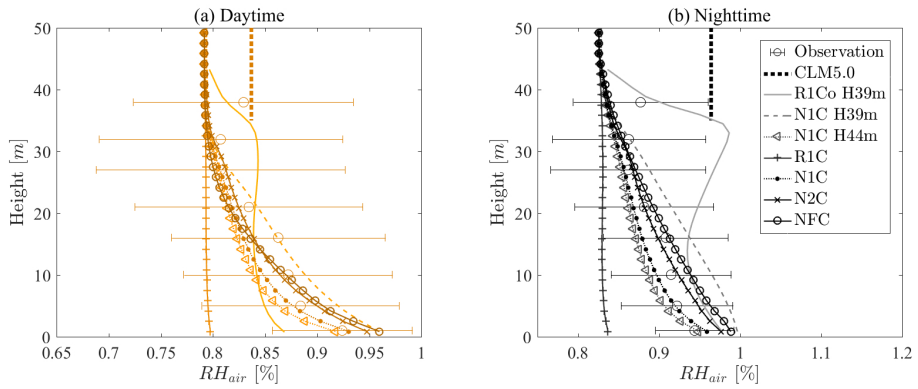

Figure 8. Relative humidity (RH) profiles as measured (open circles with error bars) and modeled (lines) during the (a) daytime and (b) nighttime. CLM5.0 is a single layer model so the profile is a single value, which refers to the relative humidity below the top of the canopy (35 m).

Night time improvements due to the multi-layered scheme could be also detected in the diurnal variation of leaf temperature [Figure 10]. The overall temperature was increased but the amplitude of leaf temperature, as well as air temperature [Figure 9], was reduced by up to 2.8 °C and the cycle followed well the observation (see N1C case).

3.5 CO_2, H_2O, Storage Fluxes and Transpiration

For CO_2 and H_2O fluxes, both H0 (full flux including the emergent tree) and H1 (partial flux) types of interpretation did not yield a successful comparison between EC data at 33 m and simulated fluxes, but the H2 type (between H0 and H2) was more valid. The EC data possibly represents a height higher than 33 m (H1) and lower than the total maximum canopy height (H0). For the H0 case, comparing to the observation (circle), the multi-layer model overestimated the total flux during the daytime in all the simulated cases [Table 2] [Figure 12]. If we assumed H1 is correct, the simulated fluxes were mostly underestimated. There were some exceptions such as a simulation (N1C H35m)

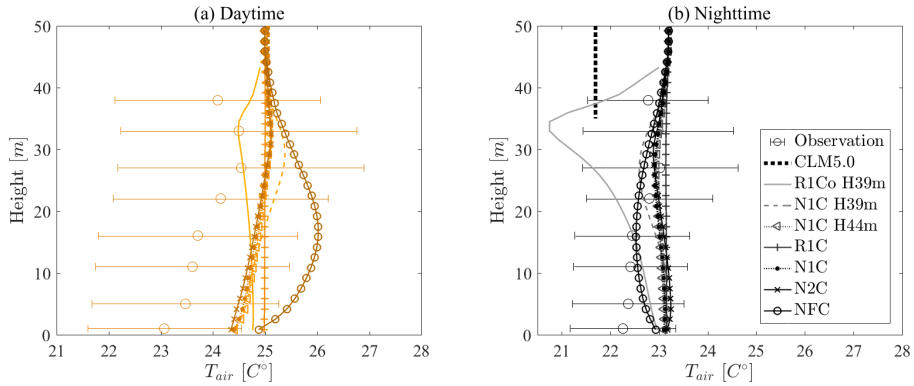

Figure 9. Air temperature profiles as measured (open circles with error bars) and modeled (lines) during the (a) daytime and (b) nighttime. CLM5.0 is a single layer model so the profile is a single value, which refers to the air temperature below the top of the canopy (35 m).

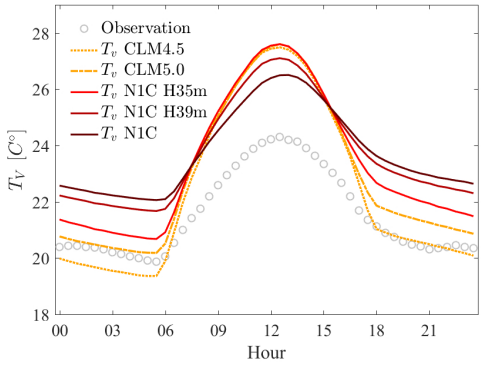

Figure 10. Diel trends in leaf surface temperature (T_V) as measured by infrared thermometry and modeled in CLM and CLM-ml. For modeled values, Eq. (7) was used to estimate the single up-scaled value from multi-layered results. The average differences between daytime and nighttime for CLM4.5, CLM5.0, N1C H35m, N1C H39m, and N1C were 5.14, 4.47, 3.97, 2.96, and 2.35 °C, respectively.

having 35 m canopy height which was a little overestimated CO_2 flux but it did not match the CO_2 profile either [Figure 7]. The other 39 m simulation (N1C H39m) and the 44 m flat canopy simulation (NFC) matched observed H_2O fluxes at 33 m but their air temperature profiles and leaf wetness were not well predicted. Therefore, both hypotheses were rejected, and the 33 m observation (eddy-covariance) would represent a height between the two places (H2 case). These results were also supported via transpiration (TR) observations, where total sap-flow rates were higher than the eddy-covariance data and lower than the simulated TR rate (Aparecido et al., 2016).

Heat storage flux was also influenced by in-canopy variability and has an important role in determining carbon and vapor fluxes. The simulated total CO_2 and H_2O fluxes with flat canopy (NFC) were relatively low [Figure 11a; Figure 12; Figure 13]. Conversely, the storage flux with NFC tended to be high in the daytime [Figure 11b]. These results indicated that precise prediction of this sensible heat storage flux is essential to estimate other fluxes and that the canopy shape profoundly influences the storage flux. This comparison also proved that that the RSL model is insensitive to in-canopy variation (R1C

vs R2C), the results were identical $R^2 = 1$, see a diurnal variation plot [Figure 11b]. However, FOC (N1C, N2C, vs NFC) showed different results in storage fluxes [Figure 11b], R^2 of N1C was 0.96 with N2C, and 0.46 with NFC.

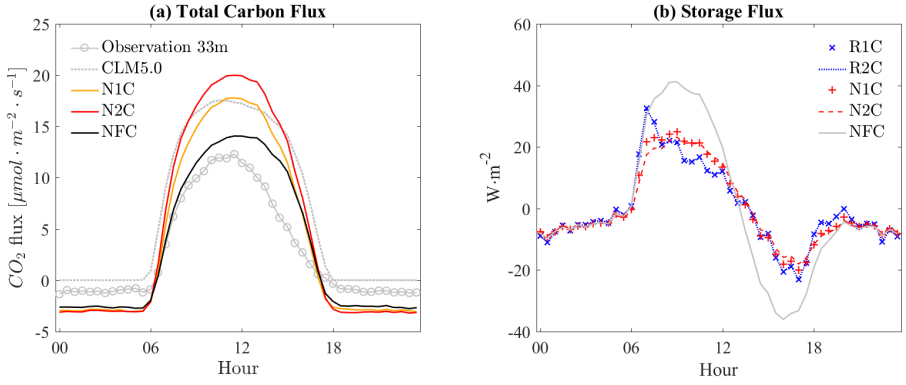

Figure 11. Diel variations in (a) CO_2 flux as affected by LAD; and (b) sensible heat storage flux as affected by turbulence scheme and LAD.

The multi-layered model showed higher night time evaporation rates, which was caused by improved turbulence scheme by model itself (RSL) but also by model's update and different LAD. These changes consequently contributed to the improvement of leaf wetness compared to the single-layered model. The nighttime water flux mostly occurred at the dense top canopy [Figure 13]. However, the comparison problem for the water flux existed similarly to CO_2 flux when compared with EC data, which made diagnosis difficult. At this time, issues with in-canopy behavior of the model can only be diagnosed through the aid of their concentration profiles. For more accurate partitioning of the water flux, all models related to canopy water need to be re-investigated with more observations from various sites to capture the effects mainly produced by the topographical complexity.

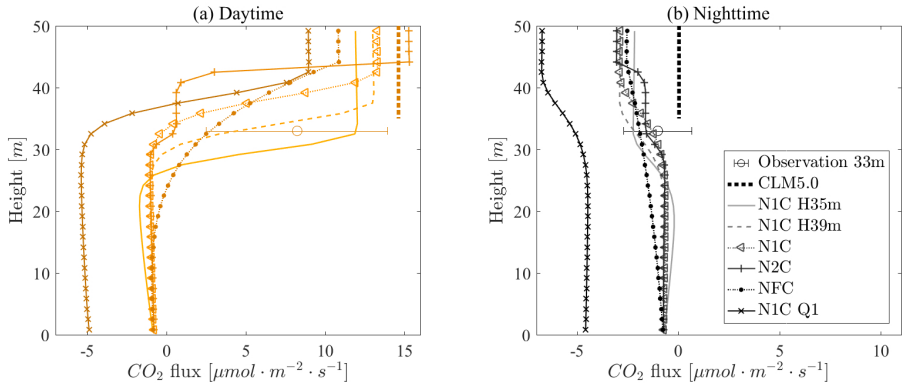

Figure 12. CO_2 flux profiles simulated using different canopy parameter sets [Table 2]. The mean of observed values is indicated by a circles and standard deviation by the bars. CLM5.0 predicts flux as a single value above the canopy, so it is shown as a constant line.

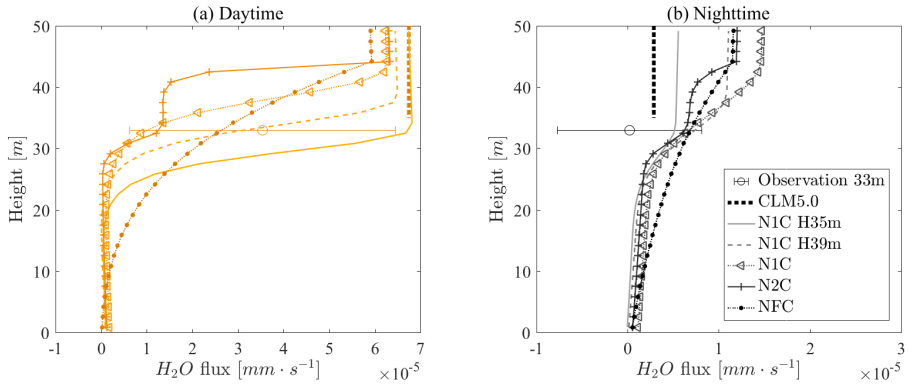

Figure 13. H_2O flux profiles simulated using different canopy parameter sets [Table 2]. For the EC measurements, mean values are represented as circles and standard deviations are indicated by bars. CLM5.0 results show the flux as a single value above the canopy.

The profile of the TR rates reasonably mimicked observations but the predicted total fluxes were still overestimated [Figure 14], 0.6-1.1 mm/day higher compared to the observation. This result supported the idea of the previous study (Song et al., 2020) which highlighted the possible issues with light-limited photosynthesis models. In this comparison analysis, the sap-flow data could be regarded as profile data, as it was measured from trees with a range of canopy heights. Aparecido et al. (2016) classified these data into three different categories: Sub-story, Mid-story, and Dominant canopy. Each tree was measured separately, and the fluxes from every tree in a category was aggregated. Also, the fluxes were accumulated (e.g., TranDom = Sub-story + Mid-story + Dominant canopy). The heights were estimated as $H_{sap} = H + \Delta H(1-r)$, following dual-canopy concept. The previous sap-flow study gave possible each maximum canopy height as H = 11 m, 27 m, and 39 m. Other parameters for H_{sap} were ΔH = 13.5m and r = 0.65, which resulted in 15.725 m, 31.72 5m, and 43.725 m for the sub-story, mid-story, and dominant canopy. These parameters were estimated through the measured data and dual-canopy model B4.

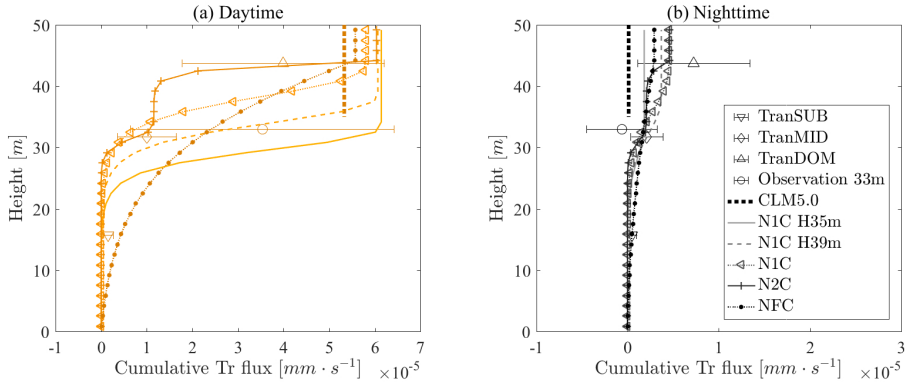

Figure 14. Transpiration (TR) profiles simulated using different canopy parameter sets [Table 2]. The observation data at 33 m is derived from the EC system and represents total H_2O flux, including TR and leaf evaporation; means values are indicated by circles. The observations with triangles at 16 m, 32 m, and 43 were derived from sap-flow data, as partitioned by the process described in Aparecido et al. (2016). Bars show standard deviations in measurements.

These simulated carbon/water/storage fluxes as well as their concentrations indicated that as compared with the single layer model, updating sub-canopy structure produced significantly different results. The model was also sensitive to the in-canopy parameters relating to the turbulence model and the light penetration model. For instance, first, GPP and transpiration (TR) predictions were mainly influenced by radiative transfer due to photosynthesis [Figure 14], which was mainly affected by LAD distribution. Then, CO_2/H_2O fluxes and their concentrations were further altered by changing turbulence-related formulations or parameters (e.g., canopy height, LAD distribution) [Figure 12; Figure 13] (also see supplementary charts). Additionally, the sensible heat storage term was also affected by LAD distributions as shown through [Figure 11]. These intrinsic variations, caused by canopy shape, produced diverse land-atmosphere interactions. We note that CO_2/H_2O fluxes, their concentrations, and GPP/Transpiration are very different terms. The fluxes represent final fluxes toward the atmosphere/in-canopy air which are affected by both the turbulence model and the underlying scalar concentrations within the canopy. GPP/TR indicates source and sink through physiological activity.

4 Summary and Conclusions

In this study, we updated and tested multi-layered CLM (CLM-ml) (Bonan et al., 2018) against a suite of micrometeorlogical observations at a tropical montane rainforest site, exploring both a multi-layered land surface model and a tropical wet forest site located on complex terrain.

The key development of the model for complex terrain was to reflect vertical leaf distribution, also known as leaf area density (LAD). This change allows for the proper representation of the spatially overlapping tree canopies at the measurement tower. Compared to simpler models such as the original CLM-ml and single-layer CLM, the CLM-ml combined with the LAD profile may improve the model's performance and give more variety of results according to a different in-canopy structure. The new LAD profile scheme was introduced based on mixed-beta distribution Eq. (2) and several shapes were applied as case studies [Figure 2].

A turbulent transfer model, a first-order closure scheme Eq. (4), was also added in the CLM-ml in order to capture the effects of varying the LAD. With the first-order closure scheme, the wind forcing data, located at the middle height of the canopy profile, were successfully included in the model [Figure 5]. Through the multi-layer scheme, this modification showed the possibility to overcome the difficulty of data usage or modeling at an complex site. It also highlighted that normal forcing data (such as Met Tower) cannot be easily applied for such steep terrain, because the wind profile abruptly increased (sensitive) at a higher elevation of canopy [Figure 5]. Additionally, a formulation for predicting in-canopy CO_2 concentrations, which includes the effects of soil respiration, was also applied to CLM-ml to allow comparisons with our available data set. This study introduced these new sub-models and described, in depth, their formulations and how to apply them to CLM-ml.

Model tests including original or updated versions exhibited that the multi-layered land model, CLM-ml, could predict variables in-depth and reflect the features of terrain better than single-layered CLM, but it requires appropriate updates to the LAD and turbulence schemes to apply more complexity. For instance, the MOST or RSL in CLM-ml was too simplified and even the RSL model which is an updated version of MOST only considered canopy height and total LAI. The surface of the study site was steep, and accordingly, overall LAD distribution became complicated, best represented by a mixed-distribution, or dual-canopy, leaf area profile [Figure 1]. Also, the displacement height needed to be re-derived based on the centroid drag on the canopy area, which resulted in the lower displacement heights by the slope effect.We need to note that the multi-layer scheme has an advantage compared to the single-layer intrinsically but also these model

updates additionally change model results throughout all important micrometeorological variables such as leaf temperature and vapor/carbon fluxes.

Simulation tests showed that the top-of-canopy energy balances predicted by CLM-ml were similar to those from CLM due to the same radiative transfer model. However, CLM-ml's main advantage was to reproduce trace gas concentrations and micrometeorological variables, allowing for partitioning of trace gas fluxes as a function of height in the under-story. Small alterations in the net radiation originated from a different energy partitioning caused by the change of model structure. The simulated meteorological profiles of variables such as air temperature, RH and CO_2 concentration were not perfectly matched with the observations. However, the mean predicted profiles fell within one standard deviation of the observations in most of the test cases in original and updated CLM-ml. We also found the CO_2 concentration profile and related fluxes were very sensitive to the soil respiration, which indicates the need for additional soil flux measurement at tropical sites for further investigation.

Applying different types of wind speed models and their parameters clearly affected the wind profile distribution. In particular, the canopy height was the most influential parameter controlling overall performance including energy, carbon, and vapor fluxes. However, changes between numerical and RSL methods and different displacement heights did not cause significant differences in the other variables (e.g., temperature, etc.), especially with the single-canopy LAD profile shape. A small difference could be detected for leaf wetness and H_2O flux when dual-canopy LAD was applied (see supplementary data). This similarity is expected, because RSL model is derived based on the first-order closure model. In contrast, some small differences indicates there is still the possibility of having a stronger influence in a different environment. Since we used only a few leaf profiles and a fixed LAI, different ambient conditions may show different outcomes. More investigation is necessary.

This study highlights that the proper selection of a turbulence model, radiative transfer model, and LAD distribution are all important for ecosystem simulation with highly diverse and complex canopy structures. For instance, LAD distributions affected the direct/instant energy exchange over the leaf profiles, so different LAD can produce significantly different outcomes for trace gas and momentum fluxes as well as state variables (e.g., CO_2/H_2O flux, sensible heat storage, GPP, TR, leaf wetness, and leaf temperature). Here, these fluxes and variables were directly controlled by the source and sink at leaves, and so are affected by light penetration. In particular, the influence of leaf wetness is more critical in this site compare to temperate forests. Simultaneously, they were also influenced by sensible heat storage flux term and turbulent transfer rate on the wind profile, which in turn was affected by in-canopy features (canopy height, LAD distribution). Hence, the turbulence model and its interaction with LAD were likewise important. The micrometeorological profiles in the air (e.g., air temperature, RH, and CO_2 concentration) were not solely/directly affected by LAD distributions (e.g., bell shape) unless the canopy shape was too different, such as a flat canopy (NFC). Rather, they are more affected by parameters for the turbulence model such as the canopy height.

This study emphasized the importance of layered structures in capturing the behavior of complex surface systems and supported previous multi-layered model studies (Jiménez et al., 2011; Ogée et al., 2003; Pitman et al., 2009; Bonan et al., 2014). Moreover, it showed the possibility of further improvement by applying a more realistic canopy shape. This multi-layered model relieves the limitations of simplified model and allows more variability caused by the internal canopy structures (Ryder et al., 2016). For example, the vertical segmentation of vapor transfer as a function of height gave a notable improvement in leaf wetness predictions. Also, GPP and TR showed that different canopy structures resulted in different total fluxes based on the vertical distribution of sources and sinks. Simulated total GPP and TR, which was computed by summing sources and sinks across all the vertical segments, was increased. However, the flux at the top toward

the atmosphere was reduced by a multi-layered turbulent scheme and by the inclusion of sensible heat storage term. These features could not be captured using a single-layer model. Updating the MOST model into RSL improved the diurnal amplitude of both leaf temperature and leaf wetness, which was too high in CLM4.5/CLM5. They were reduced by 67% and 47%, respectively, with an 470% increase in nighttime evapotranspiration in this study.

Perhaps more importantly, the micrometeorological profile measurement system provided additional, more detailed information to diagnose the model performance than is normally possible through the single-layer model. It allowed us to see the vertical variability in these variables, rather than only examining their values at the top of the canopy. Some simulations results matched fluxes at the top of the canopy, but did not reproduce the associated variables within it. For example, N1C H39 well estimated H_2O at 33 m but failed to predict leaf wetness and CO_2 profiles compared to the observations.

In conclusion, the multi-layered CLM produced intrinsic improvements over the standard version of CLM, as it was able to replicate complicated structures within the canopy. The model allowed for a much broader comparison of variables, beyond the standard average or total micrometeorological values. More importantly, using a multi-layered model resulted in significant improvement in predictions of leaf wetness, air and leaf temperatures, and relative humidity. This study suggests that LSMs could be further improved by including more detailed interpretations of the natural in-canopy phenomena, and by parameterizing using spatially rich data. We also found that model results were sensitive to canopy related parameters, especially canopy height. Among the cases in Table 2, using a numerical scheme of the FOC model, 'N - -', is recommended since it reflects LAD distributions. Also, using our target LAD distributions, '- 1C' and '- 2C', showed better results, and the complex shape, '- 1C', provided slightly better outcomes, because these showed great improvement in the surface temperature but also in the leaf wetness in this tropical site. However, as compared to the '- 1C', '- 2C' still follows the TR and PAR profiles well.

However, the multi-layered CLM still cannot sufficiently predict key micrometeorological variables. Although there is uncertainty in EC data and LAD profiles due to the site complexity, TR data still support the conclusion that CLMs normally overestimates carbon and vapor fluxes. Both the radiative transfer sub-model and the leaf distribution with height only represent a single-point horizontally , so the model cannot fully represent the effects of adjacent trees on the sloped surface. Certainly, more data-rich sites with a profile system are necessary for fairer assessment and a strong conclusion. Both vertical and horizontal profile data are required to understand spatial variability, which might be available at only a few sites.

Acknowledgments

This research was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research (Grant DE-SC0010654), and by the Korea Meteorological Administration Research and Development Program "Development and Assessment of IPCC AR6 Climate Change Scenario" (Grant KMA2018-00321). The researchers would like to thank the TAMU Soltis Center staff, particularly Dr. Eugenio Gonzales, Johan Rodriguez, Noylin Rodriguez, and Ronald Vargas Castro, for their logistics and infrastructure support. The model source codes and input files are available through `http://doi.org/10.5281/zenodo.3958251` and `http://doi.org/10.5281/zenodo.3958248`.

References

Andrews, R. S. (2016). *The temporal variation of vertical micrometeorological profiles in a lower montane tropical forest.* (Master's thesis, Texas A&M University). doi: 1969.1/157148

Aparecido, L. M. T., Miller, G. R., Cahill, A. T., & Moore, G. W. (2016). Comparison of tree transpiration under wet and dry canopy conditions in a Costa Rican premontane tropical forest. *Hydrological Processes*, *30*(26), 5000-5011. doi: 10.1002/hyp.10960

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., ... Wofsy, S. (2001, 11/01; 2017/03). FLUXNET: a new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415-2434. doi: 10.1175/1520-0477(2001)082⟨2415:FANTTS⟩2.3.CO;2

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., ... Papale, D. (2010, American Association for the Advancement of Science). Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, *329*(5993), 834-838. doi: 10.1126/science.1184984

Bonan, G. B. (2008, American Association for the Advancement of Science). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science*, *320*(5882), 1444-1449. doi: 10.1126/science.1155121

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., ... Swenson, S. C. (2011). Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research: Biogeosciences*, *116*(G2). doi: 10.1029/2010JG001593

Bonan, G. B., Oleson, K. W., Fisher, R. A., Lasslop, G., & Reichstein, M. (2012). Reconciling leaf physiological traits and canopy flux data: use of the TRY and FLUXNET databases in the Community Land Model version 4. *Journal of Geophysical Research: Biogeosciences*, *117*(G2). doi: 10.1029/2011JG001913

Bonan, G. B., Patton, E. G., Harman, I. N., Oleson, K. W., Finnigan, J. J., Lu, Y., & Burakowski, E. A. (2018). Modeling canopy-induced turbulence in the earth system: a unified parameterization of turbulent exchange within plant canopies and the roughness sublayer (CLM-ml v0). *Geoscientific Model Development*, *11*(4), 1467-1496. doi: 10.5194/gmd-11-1467-2018

Bonan, G. B., Williams, M., Fisher, R. A., & Oleson, K. W. (2014). Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum. *Geoscientific Model Development*, *7*(5), 2193-2222. doi: 10.5194/gmd-7-2193-2014

Burba, G. (2013). *Eddy covariance method for scientific, industrial, agricultural and regulatory applications: A field book on measuring ecosystem gas exchange and areal emission rates.* LI-Cor Biosciences.

Burns, S., Swenson, S., Wieder, W., Lawrence, D., Bonan, G. B., Knowles, J., & Blanken, P. (2018). A comparison of the diel cycle of modeled and measured latent heat flux during the warm season in a Colorado subalpine forest. *Journal of Advances in Modeling Earth Systems*, *10*(3), 617-651. doi: 10.1002/2017MS001248

Cai, X., Yang, Z.-L., Xia, Y., Huang, M., Wei, H., Leung, L. R., & Ek, M. B. (2014). Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed. *Journal of Geophysical Research: Atmospheres*, *119*(24), 13,751-13,770. doi: 10.1002/2014JD022113

Campbell, G. S., & Norman, J. M. (2012). *An introduction to environmental biophysics.* Springer New York.

Choudhury, B. J., & DiGirolamo, N. E. (1998, March 1998). A biophysical process-based estimate of global land surface evaporation using satellite and ancillary data I. model description and comparison with observations. *Journal of Hy-*

drology, *205*(3), 164-185. doi: 10.1016/S0022-1694(97)00147-9

Dai, Y., Dickinson, R. E., & Wang, Y.-P. (2004, 06/01; 2017/03). A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, *17*(12), 2281-2299. doi: 10.1175/1520-0442(2004)017⟨2281:ATMFCT⟩2.0.CO;2

Drewry, D. T., Kumar, P., Long, S., Bernacchi, C., Liang, X. Z., & Sivapalan, M. (2010, 12/01; 2019/03). Ecohydrological responses of dense canopies to environmental variability: 1. interplay between vertical structure and photosynthetic pathway. *Journal of Geophysical Research: Biogeosciences*, *115*. doi: 10.1029/2010JG001340

Fisher, J. B., Malhi, Y., Bonal, D., Rocha, H. R. D., Araujo, A. C. D., Gamo, M., ... Randow, C. V. (2009). The land–atmosphere water flux in the tropics. *Global Change Biology*, *15*(11), 2694-2714. doi: 10.1111/j.1365-2486.2008.01813.x

Flerchinger, G. N. (2000). *The simultaneous heat and water (SHAW) model: Technical documentation.* Northwest Watershed Research Center USDA Agricultural Research Service, Boise, Idaho.

Flerchinger, G. N., Xiao, W., Sauer, T. J., & Yu, Q. (2009, December 2009). Simulation of within-canopy radiation exchange. *NJAS - Wageningen Journal of Life Sciences*, *57*(1), 5-15. (ID: 278666) doi: https://doi.org/10.1016/j.njas.2009.07.004

Gu, L., Shugart, H. H., Fuentes, J. D., Black, T. A., & Shewchuk, S. R. (1999, 20 April 1999). Micrometeorology, biophysical exchanges and nee decomposition in a two-story boreal forest — development and test of an integrated model. *Agricultural and Forest Meteorology*, *94*(2), 123-148. (ID: 271723) doi: https://doi.org/10.1016/S0168-1923(99)00006-4

Harman, I. N., & Finnigan, J. J. (2007, 05/01). A simple unified theory for flow in the canopy and roughness sublayer. *Boundary-Layer Meteorology*, *123*(2), 339-363. (ID: Harman2007) doi: 10.1007/s10546-006-9145-6

Harman, I. N., & Finnigan, J. J. (2008, 12/01). Scalar concentration profiles in the canopy and roughness sublayer. *Boundary-Layer Meteorology*, *129*(3), 323-351. (ID: Harman2008) doi: 10.1007/s10546-008-9328-4

Hasler, N., & Avissar, R. (2007, 06/01; 2017/03). What controls evapotranspiration in the amazon basin? *Journal of Hydrometeorology*, *8*(3), 380-395. doi: 10.1175/JHM587.1

Huntingford, C., Zelazowski, P., Galbraith, D., Mercado, L. M., Sitch, S., Fisher, R., ... Cox, P. M. (2013, print). Simulated resilience of tropical rainforests to CO_2-induced climate change. *Nature Geoscience*, *6*(4), 268-273. doi: 10.1038/ngeo1741

Jiménez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., ... Wang, K. (2011, 01/27; 2019/03). Global intercomparison of 12 land surface heat flux estimates. *Journal of Geophysical Research: Atmospheres*, *116*. doi: 10.1029/2010JD014545

Katul, G. G., Mahrt, L., Poggi, D., & Sanz, C. (2004). One-and two-equation models for canopy turbulence. *Boundary-Layer Meteorology*, *113*(1), 81-109.

Kume, T., Tanaka, N., Kuraji, K., Komatsu, H., Yoshifuji, N., Saitoh, T. M., ... Kumagai, T. (2011, 15 September 2011). Ten-year evapotranspiration estimates in a Bornean tropical rainforest. *Agricultural and Forest Meteorology*, *151*(9), 1183-1192. (ID: 271723) doi: 10.1016/j.agrformet.2011.04.005

Lalic, B., Firanj, A., Mihailovic, D. T., & Podrascanin, Z. (2013). Parameterization of par vertical profile within horizontally uniform forest canopies for use in environmental modeling. *Journal of Geophysical Research: Atmospheres*, *118*(15), 8156-8165.

Launiainen, S., Katul, G. G., Kolari, P., Vesala, T., & Hari, P. (2011). Empirical and optimal stomatal controls on leaf and ecosystem level CO_2 and H_2O ex-

change rates. *Agricultural and Forest Meteorology*, *151*(12), 1672-1689. doi: 10.1016/j.agrformet.2011.07.001

Launiainen, S., Katul, G. G., Lauren, A., & Kolari, P. (2015, 24 September 2015). Coupling boreal forest CO_2, H_2O and energy flows by a vertically structured forest canopy – soil model with separate bryophyte layer. *Ecological Modelling*, *312*, 385-405. (ID: 271743) doi: 10.1016/j.ecolmodel.2015.06.007

Lawrence, D., & Vandecar, K. (2014, 12/18). Effects of tropical deforestation on climate and agriculture. *Nature Climate Change*, *5*, 27. doi: 10.1038/nclimate2430;https://www.nature.com/articles/nclimate2430#supplementary-information

Lawrence, D. M., et al. (Eds.). (2018). *Technical description of version 5.0 of the community land model (clm)* (NCAR Technical Note ed.). Boulder, CO: National Center for Atmospheric Research. (ID: 156)

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., ... Slater, A. G. (2011). Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, *3*(1). doi: 10.1029/2011MS00045

Leuning, R. (2000, 08/01). Estimation of scalar source/sink distributions in plant canopies using lagrangian dispersion analysis: Corrections for atmospheric stability and comparison with a multilayer canopy model. *Boundary-Layer Meteorology*, *96*(1), 293-314. (ID: Leuning2000) doi: 10.1023/A:1002449700617

Loescher, H. W., Gholz, H. L., Jacobs, J. M., & Oberbauer, S. F. (2005, 12/10). Energy dynamics and modeled evapotranspiration from a wet tropical forest in Costa Rica. *Journal of Hydrology*, *315*(1–4), 274-294. doi: 10.1016/j.jhydrol.2005.03.040

Maass, J., Vose, J. M., Swank, W. T., & Martínez-Yrízar, A. (1995). Seasonal changes of leaf area index (lai) in a tropical deciduous forest in west mexico. *Forest Ecology and Management*, *74*(1), 171-180.

Massman, W. J., & Weil, J. C. (1999, 04/01). An analytical one-dimensional second-order closure model of turbulence statistics and the lagrangian time scale within and above plant canopies of arbitrary structure. *Boundary-Layer Meteorology*, *91*(1), 81-107. (ID: Massman1999) doi: 10.1023/A:1001810204560

Meyers, T. P., & Paw U, K. T. (1987). Modelling the plant canopy micrometeorology with higher-order closure principles. *Agricultural and Forest Meteorology*, *41*(1), 143-163. Retrieved from `https://www.sciencedirect.com/science/article/pii/016819238790075X` doi: https://doi.org/10.1016/0168-1923(87)90075-X

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., ... Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, *116*(D12). doi: 10.1029/2010JD015139

Ogée, J., Brunet, Y., Loustau, D., Berbigier, P., & Delzon, S. (2003, 05/01; 2019/03). MuSICA, a CO_2, water and energy multilayer, multileaf pine forest model: evaluation from hourly to yearly time scales and sensitivity analysis. *Global Change Biology*, *9*(5), 697-717. doi: 10.1046/j.1365-2486.2003.00628.x

Oleson, K. W., et al. (Eds.). (2013). *Technical description of version 4.5 of the Community Land Model (CLM)* (NCAR Technical Note NCAR/TN-503 STR ed.). Boulder, CO: National Center for Atmospheric Research.

Oleson, K. W., et al. (Eds.). (2010). *Technical description of version 4.0 of the Community Land Model (CLM)* (NCAR Technical Note NCAR/TN-478+STR ed.). Boulder, CO: National Center for Atmospheric Research.

Pitman, A. J., de Noblet-Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., ... Voldoire, A. (2009, 07/01; 2019/03). Uncertainties in climate

responses to past land cover change: First results from the lucid intercomparison study. *Geophysical Research Letters*, *36*(14). doi: 10.1029/2009GL039076

Pyles, R. D., Weare, B. C., & Pawu, K. T. (2000). The ucd advanced canopy-atmosphere-soil algorithm: Comparisons with observations from different climate and vegetation regimes. *Quarterly Journal of the Royal Meteorological Society*, *126*(569), 2951-2980. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712656917` doi: https://doi.org/10.1002/qj.49712656917

Ryder, J., Polcher, J., Peylin, P., Ottlé, C., Chen, Y., van Gorsel, E., ... Luyssaert, S. (2016). A multi-layer land surface energy budget model for implicit coupling with global atmospheric simulations. *Geoscientific Model Development*, *9*(1), 223-245. doi: 10.5194/gmd-9-223-2016

Schlesinger, W. H., & Jasechko, S. (2014, 1 June 2014). Transpiration in the global water cycle. *Agricultural and Forest Meteorology*, *189*, 115-117. (ID: 271723) doi: 10.1016/j.agrformet.2014.01.011

Schlosser, C. A., & Gao, X. (2010, 08/01; 2019/03). Assessing evapotranspiration estimates from the Second Global Soil Wetness Project (GSWP-2) simulations. *Journal of Hydrometeorology*, *11*(4), 880-897. doi: 10.1175/2010JHM1203.1

Sellers, P., Berry, J., Collatz, G., Field, C., & Hall, F. (1992). Canopy reflectance, photosynthesis, and transpiration. III. a reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of Environment*, *42*(3), 187-216. doi: 10.1016/0034-4257(92)90102-P

Sheil, D. (2018, 03/20). Forests, atmospheric water and an uncertain future: the new biology of the global water cycle. *Forest Ecosystems*, *5*(1), 19. (ID: Sheil2018) doi: 10.1186/s40663-018-0138-y

Shuttleworth, W. J. (1988). Evaporation from Amazonian rainforest. *Proceedings - Royal Society.Biological Sciences*, *233*(1272), 321-346. doi: 10.1098/rspb.1988.0024

Slevin, D., Tett, S. F. B., Exbrayat, J.-F., Bloom, A. A., & Williams, M. (2017). Global evaluation of gross primary productivity in the JULES land surface model v3.4.1. *Geoscientific Model Development*, *10*(7), 2651–2670. Retrieved from `https://gmd.copernicus.org/articles/10/2651/2017/` doi: 10.5194/gmd-10-2651-2017

Song, J., Miller, G. R., Cahill, A. T., Aparecido, L. M. T., & Moore, G. W. (2020). Modeling land surface processes over a mountainous rainforest in Costa Rica using CLM4.5 and CLM5. *Geoscientific Model Development*, *13*(11), 5147-5173. Retrieved from `https://gmd.copernicus.org/articles/13/5147/2020/` doi: 10.5194/gmd-13-5147-2020

Swenson, S. C., & Lawrence, D. M. (2014). Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *119*(17), 10-10,312. doi: 10.1002/2014JD022314

Teale, N. G., Mahan, H., Bleakney, S., Berger, A., Shibley, N., Frauenfeld, O. W., ... Washington-Allen, R. (2014). Impacts of vegetation and precipitation on throughfall heterogeneity in a tropical pre-montane transitional cloud forest. *Biotropica*, *46*(6), 667-676. doi: 10.1111/btp.12166

Wang, A., Barlage, M., Zeng, X., & Draper, C. S. (2014). Comparison of land skin temperature from a land model, remote sensing, and in situ measurement. *Journal of Geophysical Research: Atmospheres*, *119*(6), 3093-3106. doi: 10.1002/2013JD021026

Wohl, E., Barros, A., Brunsell, N., Chappell, N. A., Coe, M., Giambelluca, T., ... Ogden, F. (2012, 07/15). The hydrology of the humid tropics. *Nature Climate Change*, *2*, 655. doi: 10.1038/nclimate1556

Zhang, H., Pak, B., Wang, Y. P., Zhou, X., Zhang, Y., & Zhang, L. (2013, 08). Evaluating surface water cycle simulated by the Australian Community Land

Surface Model (CABLE) across different spatial and temporal domains. *Journal of Hydrometeorology*, *14*(4), 1119-1138. Retrieved from `https://doi.org/10.1175/JHM-D-12-0123.1` doi: 10.1175/JHM-D-12-0123.1

Zhang, K., Kimball, J. S., Nemani, R. R., & Running, S. W. (2010). A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006. *Water Resources Research*, *46*(9). doi: 10.1029/2009WR008800

Zhao, W., & Qualls, R. J. (2005, 08/01; 2019/04). A multiple-layer canopy scattering model to simulate shortwave radiation distribution within a homogeneous plant canopy. *Water Resources Research*, *41*(8). doi: 10.1029/2005WR004016

Appendix A Additional Description of Forcing Data

Some forcing variables were simply taken from the top sensors of the Canopy Tower due to their low variability for the higher location. Solar data were mostly from the 44 m sensor at the Canopy Tower in the mountain forest. Air temperature, RH, and pressure were from the 38 m sensor at the Canopy Tower. The air temperature and RH did not significantly vary in the vertical direction (i.e., from 33 m - 38 m), and these data were also similar to the observation from Met Tower.

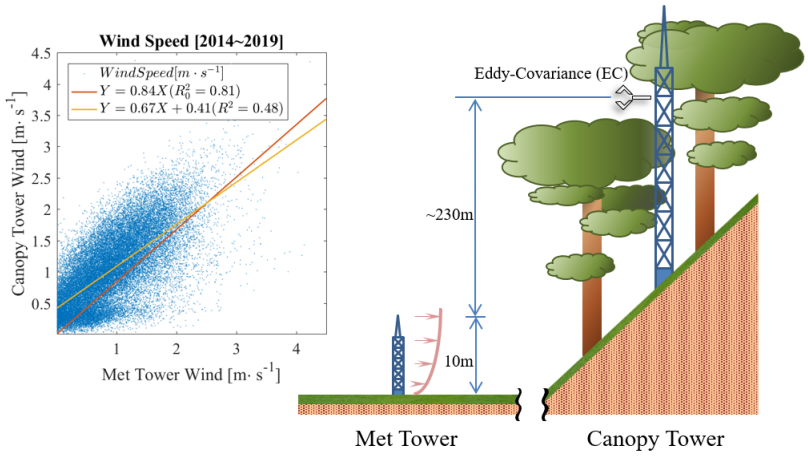

Figure A1. (a) The relationship between wind speed data as measured in the clearing at the Met Tower ($\approx$ 10 m above the ground surface) and as measured above the forest from the Canopy Tower ($\approx$ 33 m above the ground surface). The coefficient of determination (R_0^2) is calculated based on the zero interception, $Y_0 = 0$. CLM proscribes a threshold wind speed of 1 $m^2 \cdot s^{-1}$; values below that are considered to be zero to avoid numerical error (D. M. Lawrence et al., 2018). (b) Schematic of the anemometers in relation to each other, the topography, and the canopy. Predominant winds are shown into the page.

Normally, wind data measured at 10-m height over grass (Met Tower) could be applied as forcing value, if we assumed that wind log-profile of a mountain is the same as a flat surface. Moreover, 10-m is near-inertial sublayer which means wind speed is not much different at the higher location (approximately 1.6 times higher at 250 m higher in neutral condition based on the parameter in CLM5) [Figure A1].

However, Canopy Tower wind speed data were used for this study in place of those from the Canopy Tower, as it is more representative of the target location (in the forest). The vertical wind profile can differ significantly between the two places, due to the tree cover and the topography, although both observations have the same magnitude and they are reasonably correlated (R_0^2 =0.81 without interception and R^2 =0.48) as much as they provide similar results in the model [Figure A1].

Appendix B Additional Description of Wind Profile Models, CO_2 Profile Models, and LAD models

B1 Wind Speed Profile Model: First-Order Closure Model

First-order closure model is solved using a numerical method. The model Eq. (4) follows other literature (Launiainen et al., 2011; Katul et al., 2004; Drewry et al., 2010), which has

$$K_m \frac{\partial^2 U}{\partial z^2} + \frac{\partial K_m}{\partial z}\frac{\partial U}{\partial z} - C_d a(z) U^2 = 0 \tag{B1}$$

where K_m $[m^2 \cdot s^{-1}]$ is the eddy diffusivity for momentum, u $[m^2 \cdot s^{-1}]$ is wind speed, z is height, C_d is drag coefficient (0.25), and $a(z)$ $[m^2 \cdot m^{-3}]$ is LAD (Launiainen et al., 2011).

To solve the first-order closure model, the second-order derivative of wind speed with the location of the vertical grid i can be written in a numerical form as

$$\frac{\partial^2 u}{\partial z^2} = \frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta z^2} \tag{B2}$$

and the first-order derivative is

$$\frac{\partial u}{\partial z} = \frac{u_{i-1} - u_{i+1}}{2\Delta z} \tag{B3}$$

The first-order derivative for the eddy diffusivity for momentum (K_m,i) is

$$\left|\frac{\partial u}{\partial z}\right|_i = \left|\frac{u_{i-1} - u_i}{\Delta z}\right| \tag{B4}$$

The eddy diffusivity for momentum is

$$K_{m,i} = l_{m,i}^2 \left|\frac{\partial u}{\partial z}\right|_i \tag{B5}$$

where l is mixing length. Following Launiainen et al. (2011) and Katul et al. (2004), it is as

$$l_m = \begin{cases} k_v z, & z < \alpha' H_{max}/k_v \\ \alpha' H_{max}, & \alpha' H_{max}/k_v \le z < H_{max} \\ k_v(z-d), & H_{max} \le z \end{cases} \tag{B6}$$

where k_v is Von Karman constant, α' is k_v(1–d/H_{max}), H_{max} is canopy height, and d is the zero-plane displacement height (Launiainen et al., 2011; Katul et al., 2004). The displacement height is usually assumed to be $0.667 \cdot H_{max}$ but it varies in CLM-ml (Bonan et al., 2018). Then, the derivative of the eddy diffusivity is

$$\frac{\partial K_{m,i}}{\partial z} = \frac{K_{m,i} - K_{m,i-1}}{\Delta z} \tag{B7}$$

The boundary conditions for B1 are developed in two different ways in this study. In the first case, the wind speeds at two locations are known, which is appropriate for values of z below the sonic anemometer data, the EC system. In that case, the Tridiagonal matrix solution is used to solve the equation as below, and the solution provides the vertical gradient as well. Second, if one wind speed and its vertical gradient are known, the formula becomes a simple ordinary differential equations (ODEs). This case is appropriate for computing the wind profile above the sonic anemometer location, and the equation can be solved using various well-known methods such as the Midpoint Method, which we use in this study.

For the solution at points below the sonic anemometer location, applying all the numerical forms in the first-order closure model, the final relationship becomes:

$$\begin{aligned} K_{m,i}u_{i-1} - (K_{m,i} + K_{m,i+1})u_i + K_{m,i+1}u_{i+1} + \\ 0.5 \cdot (K_{m,i} - K_{m,i-1}) \cdot (u_{i-1} - u_{i+1}) - \\ C_d a(z) u_i u_i \cdot \Delta z^2 = 0 \end{aligned} \tag{B8}$$

Additionally, its matrix form to solve for Tridiagonal matrix solution becomes

$$\begin{aligned}&[K_{m,i}+0.5\cdot(K_{m,i}-K_{m,i+1})]\cdot u_{i-1}\\ -&[(K_{m,i}+K_{m,i+1})+C_d a(z)u_i\cdot\Delta z^2]\cdot u_i\\ +&[K_{m,i+1}-0.5\cdot(K_{m,i}-K_{m,i+1})]\cdot u_{i+1}=0\end{aligned} \tag{B9}$$

One may suggest using the conductivity $g_a = \rho K_m \cdot \Delta z^{-1}$ at the ground $g_{a,0}$, estimated by MOST or RSL in CLM. Usage of the conductivity was also attempted to apply as a lower boundary condition through K_m. However, it constantly produces negative wind speed at the near ground. Therefore, $g_{a,0}$ was not applied for the first-order closure model.

B2 Wind Speed Profile Model: Roughness Sublayer (RSL) Model

Calculating the wind profile using the RSL model is straightforward because it is an analytical solution. For $z > H_{max}$, it is given as:

$$\begin{aligned}u(z)=\frac{u^*}{k_v}\Bigg[ln\left(\frac{z-d}{H_{max}-d}\right)-\psi_m\left(\frac{z-d}{L_{MO}}\right)+\psi_m\left(\frac{H_{max}-d}{L_{MO}}\right)+\\ \hat{\psi_m}\left(\frac{z-d}{L_{MO}},\frac{z-d}{l_M/\beta}\right)-\hat{\psi_m}\left(\frac{H_{max}-d}{L_{MO}},\frac{H_{max}-d}{l_M/\beta}\right)+\frac{k_v}{\beta}\Bigg]\end{aligned} \tag{B10}$$

where L_{MO} is the Obukhov length [m], l_M is the mixing length [m] estimated through $l_M = 2\beta^3/(C_d \cdot a)$ which is different from the one l_m in the first-order closure model, a is the leaf area density $[m^2 \cdot m^{-3}]$ obtained via $a = LAI/H_{max}$, $u*$ is the friction velocity $[m \cdot s^{-1}]$, ψ_m is the similarity function to adjust the log profile, $\hat{\psi}_m$ is the adjusted function to accounts for canopy effects, and β is the parameter which is $\beta = u*/u(H_{max})$ (Bonan et al., 2018; Harman & Finnigan, 2008, 2007). When $z \leq H_{max}$, it is given as:

$$u(z) = u(h_{max})\exp\left[\frac{z-H_{max}}{l_M/\beta}\right] \tag{B11}$$

Finally, the wind speeds at different heights are estimated using the referenced wind speed u_{ref} at the reference location z_{ref} as:

$$u(z) = u_{ref}(z_{ref})\cdot f(z_{ref})/f(z) \tag{B12}$$

where $f()$ is a part of RSL function Eq.(B12), which includes log profile function $ln()$, similarity function $\psi_m()$, and additional adjust function $\hat{\psi_m}()$. Friction velocity $u*$ and von Karman constant k will be canceled out so they are not included in $f()$.

B3 CO_2 Profile Model

The method to estimate CO_2 profile is similar to other micrometeorological profile estimations like as RH or air temperature, described by Bonan et al. (2018). Eq. (5) is

$$\rho_m\frac{\partial C}{\partial t}-\frac{\partial}{\partial z}\left(\rho_m K_c(z)\frac{\partial C}{\partial z}\right)=\left[f_{c,sun}(z)f_{sun}(z)+f_{c,sha}(z)f_{sha}(z)\right]a(z) \tag{B13}$$

where ρ_m is molar density $(mol \cdot m^{-3})$, C is CO_2 concentration $(\mu mol \cdot m^{-2}s^{-1})$, t is the temporal space, K_c is scalar diffusivity which is the same as K_m in this study due to the ratio is close to 1 (Launiainen et al., 2011), f_{sun} is the fraction of sunlit leaves, f_c is photosynthesis flux, and the sum of square bracket is source and sink term (f_c). In a numerical form with vertical grid index i and temporal grid index t, it can be written as

$$\begin{aligned}\rho_m\frac{\Delta z}{\Delta t}\left(C_i^{t+1}-C_i^t\right)-g_{a,i-1}\left(C_{i-1}^{t+1}-C_i^{t+1}\right)+g_{a,i}\left(C_{i-1}^{t+1}-C_{i+1}^{t+1}\right)\\ =f_{c,i}^{t+1}L_i(z)\end{aligned} \tag{B14}$$

where L leaf area index at each height which is estimated by $a(z)\cdot\Delta z$, and g_a is aerodynamic conductance $[mol\cdot m^{-2}s^{-1}]$ which estimated through $\rho_m\cdot K_c/\Delta z$. Like as the first-order closure model, its matrix form to solve for Tridiagonal matrix solution becomes

$$\begin{aligned} -g_{a,i-1}C_{i-1}^{t+1} \\ \left(\rho_m\frac{\Delta z}{\Delta t}+g_{a,i-1}+g_{a,i}\right)C_i^{t+1} \\ -g_{a,i}C_{i+1}^{t+1} \\ = f_{c,i}^{t+1}L_i(z)+\rho_m\frac{\Delta z}{\Delta t}C_i^t \end{aligned} \tag{B15}$$

B4 LAD Profile Distribution Model

As described before, Beta function is used for the single-canopy LAD distribution model as in (Bonan et al., 2018).

$$f_{LAD,1}(z,H)=\frac{L}{H}\cdot f_{Beta}(z/H,p,q)+\frac{S}{H} \tag{B16}$$

where $f_{LAD,1}$ $[m^2\cdot m^{-3}]$ is the single-canopy model for leaf area density (LAD), z [m] is height, H [m] is canopy height, L $[m^2\cdot m^{-2}]$ is leaf area index, S is stem area index, and p and q are shape parameters for Beta function. Then, the two-canopy (dual-canopy) LAD model can be written using mixed-distribution as

$$\begin{aligned} f_{LAD,2}(z,H_d,r,\Delta H)=r\cdot f_{LAD,2}(z+(r-1)\cdot\Delta H,H_d)+ \\ (1-r)\cdot LAD_1(z+r\cdot\Delta H,H_d) \end{aligned} \tag{B17}$$

where $f_{LAD,2}$ $[m^2\cdot m^{-3}]$ is the two-canopy model for leaf area density (LAD), r [-] is horizontal relative location between the two canopies, H_d [m] is the height of dominant tree which is the same as H, the maximum canopy height H_{max} [m] between the two is estimated through $H_{max}= H_d+(r-1)\cdot\Delta H$, and ΔH [m] is the vertical distance between each canopy heights H_d [Figure 1]. Adding more combinations for mid-story could reach four-canopy model. If mid-story and the dominant tree had the same canopy shape parameteres as $f_{LAD,1}$, the four-canopy model could be written as

$$\begin{aligned} f_{LAD,2,2}(z,H_d,r_d,\Delta H_d,H_m,r_m,\Delta H_m,v)=v\cdot f_{LAD,2}(z,H_d,r_d,\Delta H_d)+ \\ (1-v)\cdot f_{LAD,2}(z,H_m,r_m,\Delta H_m) \end{aligned} \tag{B18}$$

where $f_{LAD,2,2}$ $[m^2\cdot m^{-3}]$ is the four-canopy model: the subscript means two horizontal canopies and two vertical canopies, these parameters are the same as $f_{LAD,2}$ but subscript d represents the dominant tree and m represents the mid-level tree, and v is LAI ratio between dominant trees and mid-story.

Finally, the parameters are fitted based on measured LAD profile data (Song et al., 2020), using the least-squares method. Here, for proper fitting, all LADs from both measurement and model are converted in a cumulative form because the main purpose of fitting is for the light-extinction model. For the four-canopy model, due to too many parameters, some assumptions were made: the horizontal distribution mid-story is homogeneous and between mid-size trees have no gaps (no slope effect) which make r_m unnecessary. For the two-canopy LAD fitting, it was very unstable due to low level canopy (mid-story), so LAD below the 20m was not included while the fitting process. For the single-canopy fitting, the fitted shape were similar whether LAD below the 20m is included or not. The mean-least-squared (MLS) value was 0.066 for $f_{LAD,1}$. Bonan et al. (2018) briefly explains several LAD shapes for deciduous tree and pine tree based on Beta distribution and pine tree ($p = 11.5$ and $q = 3.5$) is more close to our observed LAD shape (MLS is 1.1). $f_{LAD,2}$ has highest MLS due to the region of the mid-story: it is because the fitting was made for above 20m but the MLS is estimated for the whole canopy levels [Table B1].

Table B1. Fitted parameters in different LAD model

ID	$f_{LAD,x,v}$	p	q	H_d	H_m	ΔH_d	ΔH_m	r_d	v	MLS
M1	$f_{LAD,1,1}$	0.9	0.4	41.5	-	12.4	-	1.00	1.00	0.0658
M2	$f_{LAD,2,1}$	69.9	8.7	39.5	-	13.5	-	0.65	1.00	0.6342
M3	$f_{LAD,2,2}$	51.5	5.7	39	10	13.7	0	0.66	0.81	0.0112

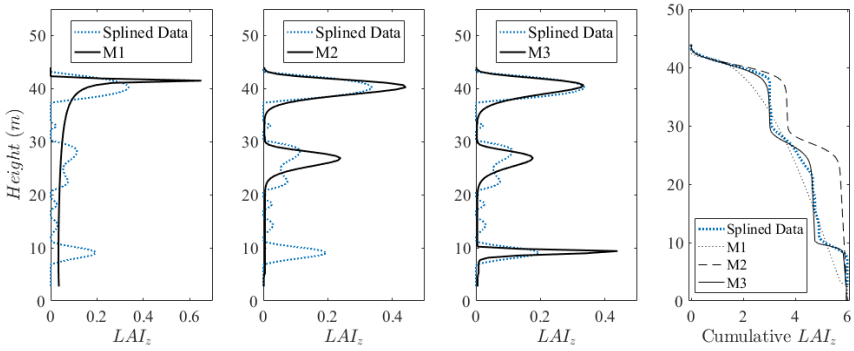

Figure B1. LAD profile test. Dotted lines refer to observation which is directly interpolated and estimated from light distribution, through light-extinction model.