IMPROVING PREDICTIONS OF SOIL AND PLANT EVAPOTRANSPIRATION IN VADOSE ZONE AND LAND SURFACE MODELS

A Dissertation

by

JAE YOUNG SONG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Anthony T. Cahill
Co-Chair of Committee,	Gretchen R. Miller
Committee Members,	Huilin Gao
	Binayak Mohanty
Head of Department,	Robin Autenrieth

December 2019

Major Subject: Civil Engineering

ABSTRACT

This study focuses on the surface hydrology to contribute to its better parameterization through inter-comparison and direct improvement with two scopes: (1) soil evaporation process, and (2) the forest hydrological/physiological process based on Community Land Model (CLM).

In soil evaporation study, the behaviors of the in-soil system and soil-atmosphere interface are explored and improved. The study of saturated front depths, under steady-state bare-soil evaporation, updates the previous method, which also provides a better understanding of the mechanism of the in-soil system. Through this investigation, the analytical and mass-conservative solution of the front depth is developed from Darcy's equation by applying two-direction flows in the soil system. More importantly, applying soil-pore heterogeneity at the soil-atmosphere interface in the model has a significant influence on the evaporation rate. Through this study, a practical equation to use soil-pore heterogeneity is selected among known diffusion-based models (except empirical models) for the above the soil layer. Also, this study verifies the fully physical-based model about in and out soil system can mimic the behavior of bare-soil evaporation.

The study of the canopy process identifies that CLM4.5&5 fail to capture the environmental complexity on tropical mountain rainforest in Costa Rica. The newer version (CLM5) shows some improvement. However, it still has discrepancies with observations. This study highlights the issue of the parameter for photosynthesis, but also the lack of in-canopy variability caused by overly simple model structure for sub-canopy layers and site-specific features (e.g., large/frequent precipitation, steep slope). Conversely, multi-layered CLM (CLM-ml) alleviates temperature-related variables and leaf wetness, and it is useful in self-diagnosing through the profiled observation. For this test, the CLM-ml is updated for an in-canopy turbulence transfer, canopy shape, and carbon dioxide (CO_2) concentration profile. This study indicates updating sub-canopy structure (e.g., canopy shapes), and the parameter for the turbulent transfer model can have a significant influence on model performance. However, more extensive monitoring of sub-canopies is necessary to increase model reliability for this and other sites with complex terrain and vegetation roughness.

DEDICATION

To the God, my family, and all who supported me, and who prayed for me with love.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisors, Professors Anthony Cahill and Gretchen Miller, who have guided and supported me through dedication, inspiration, knowledge, and creativity. Moreover, thanks for being warm and helping me to have various research experiences. Also, I would like to thank my research colleagues, Professor Georgianne Moore and Dr. Luiza Aparecido, who have worked together for the Costa Rica project. They shared their knowledge and helped me to understand the physiological process to accomplish this work.

Furthermore, I also thank my friend, Dr. Jonggun Kim, who initially helped me to handle a complex model, and Ryan Andrews and Courtney Market who helped me smoothly start the Costa Rica project. I am also thankful to my colleagues in the Miller Geoecohydrology Research Group who cheer and help each other. I am grateful to the TAMU Soltis Center staff, particularly Dr. Eugenio Gonzales, Johan Rodriguez, and Ronald Vargas Castro, for their logistics and infrastructure support. I also thank other committee members who have contributed to the successful finish of my career here at College Station: my dissertation committee members, Professors Binayak Mohanty and Huilin Gao; my exam committee member, Professor Francisco Olivera; and my numerous classroom instructors.

Moreover, I am grateful to those who have been supportive in my moments of despair, particularly my mother, father, sister, my girlfriend who would have been a more difficult time than anyone. Also, I appreciate other close people who care and love me, Doctor of Medicine Garry Gore and Kimiye Cabrera who fast diagnose the problem, Doctor of Medicine Hyeon-Seok Eom who treated it successfully, other numerous medical doctors and nurses who cheered me up, Professor Anthony Cahill and Gretchen Miller for their support and help me during the treatment, and my friends who prayed for me in love.

I also want to thank my friends and pastors at the A&M Korean Student Church. They have prayed for me and changed my life, too. I also appreciate the alumni of Hanyang University's civil engineering department. They gave me the feeling of staying here with my real family. Moreover,

I want to thank my basketball team for leaving me so many memories.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Anthony Cahill (advisor), Gretchen Miller (co-advisor), and Huilin Gao of the Department of Civil Engineering and Professor Binayak Mohanty of the Department of Biological and Agricultural Engineering.

The data measurement and analysis for Chapter 4 and Chapter5 were conducted in collaboration with Professors Gretchen Miller, Anthony Cahill, and graduate student Ryan Andrews of the Department of Civil Engineering and Professor Georgianne Moore and Dr. Luiza Aparecido of the Department of Ecosystem Science and Management.

All other work conducted for the thesis (or) dissertation was completed by the student independently.

Funding Sources

1. Research for Chapter 4 and Chapter5 was supported by the Office of Science (BER) U.S. Department of Energy (DE FOA-0000749).

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iii

ACKNOWLEDGMENTS iv

CONTRIBUTORS AND FUNDING SOURCES vi

TABLE OF CONTENTS vii

LIST OF FIGURES x

LIST OF TABLES xvi

1. INTRODUCTION 1

2. SATURATED DEPTH (CHARACTERISTIC LENGTH) OF POROUS MEDIA UNDER STEADY-STATE EVAPORATION 6

    2.1 Introduction 6
    2.2 Methodologies 11
        2.2.1 Theory 11
        2.2.2 General Formulation 15
        2.2.3 Special Case (Closed-Form) 16
    2.3 Simulation Setup 19
    2.4 Results and Discussion 21
    2.5 Summary and Conclusion 31

3. INVESTIGATION OF DIFFUSION-BASED SOIL EVAPORATION MODEL FOR HETEROGENEOUS PORE SIZES AND ITS APPLICATION WITH SOIL RESISTANCE SCHEME 33

    3.1 Introduction 33
    3.2 Two Homogeneous Diffusion Models 36
        3.2.1 S&M Model 37
        3.2.2 SCH Model 38
        3.2.3 The Proof of Similarity 38
    3.3 Diffusion Model for Heterogeneous Soil Grain 40
        3.3.1 Assumptions and Basic Equations 41
        3.3.2 Effective Pore Size 43

3.3.3 Effective Surface Water Content 44
3.4 Setup for Model Test 46
3.4.1 Modified-SCH Model 47
3.4.2 Full Resistance Scheme 47
3.4.3 Simulation Setup 49
3.5 Results and Discussion 51
3.6 Summary and Conclusion 55

4. MODELING LAND SURFACE PROCESSES OVER A MOUNTAINOUS RAINFOREST IN COSTA RICA USING CLM4.5 AND CLM5 58

4.1 Introduction 58
4.2 Methodology 62
4.2.1 Study Site 62
4.2.2 Micrometeorological Measurements 62
4.2.3 Model Description 64
4.2.4 Simulation Setup and Comparison Method 68
4.3 Simulation Results and Comparison 72
4.3.1 Net Radiation and Albedo 72
4.3.2 CO_2 Flux (GPP) 73
4.3.3 H_2O Flux 76
4.3.4 Leaf Wetness 78
4.3.5 Temperatures and Soil Flux 80
4.4 Discussion and Conclusion 86

5. PROFILE MODELING OF MICROMETEOROLOGICAL VARIABLES IN A TROPICAL MOUNTAIN RAINFOREST USING MULTILAYERED CLM (CLM-ML) 93

5.1 Introduction 93
5.2 Methodology 96
5.2.1 Micrometeorological Measurements 96
5.2.2 Forcing Data 98
5.2.3 Forcing and Canopy Heights 99
5.2.4 Simulation Setup and Analysis 100
5.2.5 Additional Calculations 105
5.3 Simulation Results and Discussion 108
5.4 Summary 118
5.5 Conclusion 121

6. SUMMARY 122

REFERENCES 126

APPENDIX A. HYPERGEOMETRIC FUNCTION 143

APPENDIX B. UPDATED CALCULATIONS FOR MULTI-LAYERED CLM 144

B.1 Additional Plots 144
B.2 Additional Description of Wind Profile Models and CO_2 Profile Models 145
B.2.1 Wind Speed Profile Model: First-Order Closure Model 145
B.2.2 Wind Speed Profile Model: Roughness Sublayer (RSL) Model 147
B.2.3 CO_2 Profile Model 148
B.2.4 LAD Profile Model 149

APPENDIX C. SUPPLEMENTAL SOURCES CONSULTED 152

LIST OF FIGURES

FIGURE Page

1.1 (a) Energy Balance, and (b) Water Balance. 2

2.1 Water supply machanism for drying soil in pore scale. Δz, Δh, and $\Delta\theta$ are the vertical difference of height, pressure head, and the water contents. q_z, q_x, and e_0 are vertical water flux, horizontal water flux, and surface evaporation flux. The Figure is modified from (Lehmann, 2008) 8

2.2 The scheme of evaporation process. With the idea that the evaporation rate occurs only at active pores (thin films), the water under the inactive pores which is called as vertically stable water (VSW) in this manuscript can migrate through vertical flux under the active pores. 12

2.3 Test Simulation using Eq. (15): c = 1 (Sadeghi et al. (2012)) and c = 2 (New method). 18

2.4 Estimated front depth ($L_c = Z_{max}$) vs VGM parameter 'n' where e_0 is 3 mm/day and 10 mm/day. Dotted lines are estimated by HYDRUS and general lines are calculated based on using previous method Eq. (2.8) with ($c = 1$). The Root Mean Squared Errors (RMSE) of the prediction in this graph are 12.114 mm for 3 mm/day and 3.151 mm for 10 mm/day. 22

2.5 Days for the end of Stage I vs VGM parameter 'n'. Lines are earned through Eq. (2.10) and Eq. (2.11) with ($c = 1$) and marks from HYDRUS. 23

2.6 Estimated front depth ($L_c = Z_{max}$) vs VGM parameter 'n' where e_0 is 3 mm/day and 10 mm/day. Dotted lines are estimated by HYDRUS and general lines are calculated based on using previous method Eq. (2.8) with ($c = 2$). The RMSE in this graph are 1.875 mm for 3 mm/day and 0.993 mm for 10 mm/day. 24

2.7 Days for the end of Stage I vs VGM parameter 'n'. Lines are earned through Eq. (2.10) and Eq. (2.11) with ($c = 2$) and marks from HYDRUS. 25

2.8 Plots of total mass loss with different soil properties for comparing results of HYDRUS and Eq. (2.10): Old version is with ($c = 1$) and new version is with ($c = 2$). . 26

2.9 Dotted lines are new method and the lines are generated by HYDRUS (e_0 = 3mm/day). The filled circles indicates the end of Stage I (h=-1500kP). 27

2.10 Detailed plots front depth behavior with different parameter 'n' in HYDRUS simulation. 28

2.11 The same graph as [Figure 2.6] but the front depth is defined at the water content equal to 0.95 in results from HYDRUS. In [Figure 2.6], it is defined as 0.99. 29

2.12 Experiment data plot using Z_{max} based on Eq. (2.18): it is the same as Sadeghi et al. (2012) if $c = 1$, and it is from the introduced method if $c = 2$. 30

3.1 Plots of S&M models and SCH model (Suzuki and Maeda, 1968; Schlünder, 1988). 34

3.2 The concept of S&M models: P is a wetted part for evaporation e, W is a width of wet and dry surface, and δ is viscous boundry layer which is usually determined by windspeed. SCH model, on the other hand, consider 2-D circle for P and 2-D square for W 37

3.3 Comparison between SCH Eq. (3.4) and modified S&M Eq. (3.5) 40

3.4 Scheme of the water behavior at a surface. Within Viscous Boundary Layer (VBL), k/k_0 is the function of $\Theta_{eff(2D)}$, P_{eff}, and δ which are the effective surface water contents, the effective pore size, and the thickness of VBL. In the soil layer, the vertical flow q_z follows Darcy's law, which consists of $\Theta_{eff(2D)}$, τ, K_s, $\partial h/\partial z$, and n, where τ is the tortuosity and n is a parameter for soil feature if the Van Genuchten (VG) retention curve (van Genuchten, 1980) is used. 41

3.5 The behavior of effective pore size based on Eq. (3.10). Mean of pore size is set to 100 μ m, the parameters n are 1.2, 1.5, 2, 3, 8 and $m = 1 - 1/n$. 45

3.6 Simulation using SCH model using the effective pore size and surface water content (lines). Average pore size 100 μm and δ=0.0023m are used here. Original SCH model is depicted for the reference (dotted line). 46

3.7 Plots in (A) are total resistance ($r_t = r_s + r_{BL}$) using Eq. (3.14). (B) is relative evaporation from Eq. (3.17) or Eq. (3.18). In plot (B), lines are simuated based on original SCH model Eq. (3.4) and dotted lines are from modified SCH Eq. (3.13) which is applied effective pore sizes and effective water contents. Marks are measured data by Aluwihare and Watanabe (2003) and Griend and Owe (1994). Additionally, fitted line by Griend and Owe (1994) also presented on the graph. 51

3.8 Plots in (A) are total resistance using Eq. (3.14). (B) is relative evaporation from Eq. (3.17) or Eq. (3.18). In plot (B), lines are simuated based on SCH model Eq. (3.4) and dotted lines are from modified SCH Eq. (3.13) which is applied effective pore sizes and effective water contents. Marks are measured data by Kondo et al. (1990). 52

3.9 a similar graph as [Figure 3.7] and [Figure 3.8]. Marks are measured data by Chanzy and Bruckler (1993). 52

3.10 lines are simulated based on original SCH model Eq. (3.4) and dotted lines are from modified SCH Eq. (3.13) which is applied effective pore sizes and effective water contents. Additionally, $\frac{\partial h}{\partial z} = 1 + \frac{e_0}{4K(h_c)}$ is applied following Lehmann et al. (2018). The datasets and parameters are produced by Lehmann et al. (2018); Merlin et al. (2016). 53

3.11 plots to see influence of effective water contents, effective pore sizes and hydraulic conductivity. Applying SCH means the resistance of boundary layer (r_{BL}) is active, and K(θ) (Hydaulic Conductivity) means soil resistance (r_s) is applied. Modified SCH means applying effective pore size and effective water content..... 54

4.1 Sketch of Canopy Tower located in a plot within a mature premontane moist tropical forest in Costa Rica (right) with LAI profiles highlighted (left) along with the location of the eddy covariance system (EC 33m) and the spire (44 m) hoisting the net radiometer including an incoming solar radiation sensor. The sum of points (LAI_z) totals $LAI(\mathrm{m}^2 \cdot \mathrm{m}^{-2}) = 6$ for this stand. The method used to obtain the LAI profile of the canopy is explained in a previous study Andrews et al. (2019)..... 64

4.2 Resistance network schemes incorporated within CLM for a) sensible heat flux and b) latent heat flux. Main state variables are atmospheric forcing temperature (θ_{atm}) and specific humidity (q_{atm}), canopy air temperature (n) and specific humidity (q_s), leaf temperature (T_v) and its corresponding specific humidity (q_v), and ground temperature (Tg) and its corresponding specific humidity (q_g). Relevant heights are the atmospheric reference height (z_{atm}), the canopy roughness heigth (Z_0), the groundwater roughness height (Z_0'), and the displacement height (d). Resistances are specified by their scalar (h for heat and w for water vapor), type (a for aerodynamic, b for boundary layer, s for stomatal, or $litter$ for litter), and lighting (sun or $shade$). Leaf wetness also exerts control on fluxes, via a wetness fraction (f_{wet}) and ($L + S$) is leaf and stem are index. Figure adapted after Oleson et al. (2013). 66

4.3 (a) and (b) - comparison of net radiation between CLM and measurement on Canopy Tower at 44m. (c) and (d) - albedo at 44m. (e) and (f) - PAR comparision for shaded canopies. All left plots (a, c, and e) are ensemble diurnal variation and the right plots (b, d, and f) are one to one comparison plots between CLM and measured data. Hysteresis depicted on (d) and (f) is based on hourly ensemble average values for daytime..... 74

4.4 (a) The ensemble diurnal variation of CO_2 flux: differences between Eddy Covariance (Canopy Tower 33m) and CLM in daytime is 6.7, 4.9, and 3.6 $\mu\mathrm{mol}\cdot\mathrm{m}^{-2}\mathrm{s}^{-1}$ for CLM4.5, CLM5, and CLM5BGC, respectively; (b) One to one plot in reference to data shown in figure (a); (c) APAR vs GPP, and w_j is simulated with default parameters and $\mathbb{C}_i$=0.667; (d) 'A)' is the theoretical maximum $\alpha\approx0.11$, 'B)' is saturated/elevated condition $\alpha\approx0.075$ (absence of photorespiration), 'C)' is normal atmosphere condition $\alpha\approx0.05$ if $\Phi\approx0.6$, and 'D)' is a fitted value $\alpha\approx0.021$ if $\Phi\approx0.25$ from our observation. [Θ] means applying the hyperbolic function Eq. (4.7) with the curvature value Θ. The change of the slope can alter maximum assimilation rate, and the alteration can be counterbalanced if Θ is modified. 76

4.5 Test simulation using Φ=0.25 and Θ=0.7 as in [Figure 4.4d]. By the modification of Φ, maximum GPP has been reduced. It is possible to improve this model by updating Θ. r^2 is a R-squared value without an intercept. 77

4.6 (a) The ensemble diurnal variation of total H_2O flux, where "Measured 33m" is measured by Eddy Covariance (at 33m), "Sapflow" is transpiration measured through sapflow, all "CLM" are about evapotranspiration (ET), 'fmx=1' sign represents fwetmax=1, and 'Φ=0.25' means 0.25 applied to Φ in Eq. (4.7); (b) One to one plot in reference to data shown on figure (a); (c) diurnal variation of partitioned H_2O flux, where ET, TR, and VE are evapotranspiration, transpiration, and canopy evaporation from CLM; (d) The one to one plots of 30 min scale transpiration rate between simulated TR versus sapflow transpiration rates; (e) The one to one plots of daily evapotranspiration (except nighttime transpiration) rate between simulated ET versus measured one from eddy-covariance; (f) The one to one plots of daily transpiration rate between simulated TR versus sapflow data (except nighttime transpiration) 82

4.7 (a) Transpiration/Evapotranspiration versus leaf wetness and classified by relative humidity [0-1]; (b) The ensemble diurnal variation of leaf wetness, where '38m', '11m', and '3m' are measurement heights and the others are leaf wetness from CLM5 (f_{wetmax}=0.05), CLM5 fmx=1 (f_{wetmax}=1), and CLM4.5 (f_{wetmax}=1); (c) and (d) are the behavior of wetting canopy at 38m and large markers are the average of each simulated value. The collected data has the condition that there is no previous rainfall event at least in 2 hours and leaf wetness is lower than 0.2 at the beginning; (c) is for 0.5-hour rainfall events (one consecutive event in 30-min scale); (d) is for 2 hours rainfall events (four consecutive events); (e) and (f) are the behavior of drying canopy. For this analysis, Consecutive no-rainfall (3 hours) data and daytime between 10 am and 2 pm are collected right after rain event. Marked lines are from measurements, and continuous lines are estimated from CLM; (e) is when solar radiation is low (0 - 300 $\mathrm{W}\cdot\mathrm{m}^{-2}$); (f) when solar radiation is higher than 300 $\mathrm{W}\cdot\mathrm{m}^{-2}$. All leaf wetness [0-1] from CLM was re-estimated based on Eq. (4) because such CLM5 has a limit as f_{wetmax}=0.05. Also, whether or not to apply the power of 2/3 does not change much our comparisons. 83

4.8 The ensemble diurnal variation of air temperatures. Canopy Levels at 22-33m and 1-5m are measured air temperatures, Ta represents air temperature at 28.075m in CLM, and Tg is ground air temperature at 0.01m in CLM. 'Ts -0.02m' is measured/simulated soil temeprature. In (a), both CLM4.5 and CLM5 is overestimated in daytime (+0.8 and +1 °C) and underestimated during the nighttime (-1.9 and -1.1 °C). In (b), differences bettwen 'Measured Ta 01-05m' and all CLM values (CLM5.0 Tg, CLM4.5 Tg, CLM5.0 Ts, and CLM4.5 Ts) are -0.39, -0.14, -0.32, and -0.06 in daytime and -0.02, 0.18, -0.11, and 0.08 °C in nighttime. Differences with 'Measured Ts -0.02m' are -0.04, 0.21, 0.03, and 0.30 in daytime and 0.90, 1.10, 0.81, and 1 °C in nighttime. 84

4.9 The ensemble diurnal variation of soil/ground heat fluxes (into soil +) (left) and soil evaporation. 84

4.10 (a) The diurnal variation of leaf temperatures with measured canopy air temperatures; (b) The one to one plot of leaf temperatures: CLM vegetation temperatures (Tv) are compared with measured values for the both gauged shade (Shade Tv) and sunlit (Sunlit Tv) vegetation temperatures; (c) The one to one plots about measured canopy air temperatures versus measured leaf temperatures (sunlit and shade); (d) The one to one plots about canopy air temperatures versus leaf temperatures from CLM (CLM5 Ta vs CLM5 Tv) and observation (Canopy Ta 22-38m vs averaged Tv from sunlit and shade Tv): Averaged Tv is estimated through $(LAI_{Shade} \times Tv_{Shade} + LAI_{Sunlit} \times Tv_{Sunlit})/LAI$. In (a), daytime differences 'CLM 5.0 Tv' minus measurments ('Measured Ta 22-38m', 'Measured Shade Tv', and 'Measured Sunlit Tv') are 1.1, 2.4, and 1.0. In nighttime, the differences are -2, -0.3, and -1.8 °C. CLM5 normally 0.2 higher in daytime and 0.8 higher in nighttime. 85

5.1 Simplified illustration of sloped surface and its two canopies for this study site, where H is average canopy height in large scale, H_{max} is the maximum canopy height at the measurement place, r is the measurement location between two canopies, EC represent eddy-covariance measurement or its sonic anemometer, and d is displacement height. 97

5.2 Simplified illustration of two different wind sensors. R_0^2 is R-squared without intercept. CLM only takes wind value higher than $1\,\mathrm{m}^2 \cdot \mathrm{s}^{-1}$ to avoid numerical error (Lawrence et al., 2018) 101

5.3 Various LAD distributions by the two parameters (p, q, r) and heights $(h = H_{max})$ for this study [Table 5.1]. 30 nodes are used here. 106

5.4 (a) and (b) - comparison of net radiation between CLM and measurement on Canopy Tower at 44m. 109

5.5 (a) PAR profile, (b) Net radiation profile, (c) Sensible heat flux, and (d) Evaporative fraction in daytime. These profiles are simulated through multi-layered CLM (CLMML) with different structures [Table 5.1]. 110

5.6 Wind profile. (a) Wind profiles with different canopy heights. H indicates H_{max}=$H(d)$ [Table 5.1]. (b) Wind profile with different models. (c) Wind profile with different LAD distributions. 111

5.7 Diurnal variation plot for canopy water which is up-scaled from profile results [Figure B.3 in Appendix B.1] via Eq. (5.3). Here the wetness is not used to avoid the confusion by the power term (2/3) in CLM. This canopy water represents the ratio between the current canopy water and the maximum canopy water $W_{max} = 0.1\mathrm{kg}\cdot\mathrm{m}^{-2}$ without the power term, because W_{max} is a constant and total LSAI has only temporal variability. The water on y-axis represents the total water on the canopy : the LSAI is approximately $\approx 6\ \mathrm{m}^2\cdot\mathrm{m}^{-2}$. 112

5.8 CO_2 concentration profile, with observations from AP200 system 113

5.9 Relative Humidity (RH) profiles with the profiled observation. 114

5.10 Air Temperature profiles. 114

5.11 CO_2 flux profiles simulated using under different canopy parameters [Table 5.1]. For the observation from EC measurement, the circle is the mean value and the bar indicates standard deviation. 116

5.12 H_2O flux profiles in the same way of [Figure 5.11]. 116

5.13 Transpiration (TR) profile in the same way of [Figure 5.11]. It should be noted that the Observation 33 m is EC which represents total H_2O flux including TR. TranSUM, TranMID, and TranDOM are from sap-flow data, which represent in 15.725 m, 31.725 m, and 43.725 m for the Sub-story, Mid-story, and Dominant canopy 117

5.14 Diurnal plot of Leaf Temperature (T_V) via Eq. (5.3). 117

B.1 Comparison plots for wind speed data between Canopy Tower and Met Tower. Slope and its correlation get higher when wind speed is higher. CLM uses wind speed bigger than $1\ \mathrm{m}\cdot\mathrm{s}^{-1}$. The fitting is only for a slope without intercept so the R-squared R_0^2 is normally higher than usual case. 144

B.2 (a) Soil respiration data from Rivera et al. (2018) and fitted parameters. (b) Temperature fitting between forcing temperature and ground temperature. 145

B.3 Leaf wetness profile. 145

B.4 LAD profile test. 151

LIST OF TABLES

TABLE Page

4.1 Fitting parameters and regression coefficients for sap flow and eddy-covariance measurements (except nighttime transpiration) versus simulation by CLM in daily scale for [Figure 4.6e] and [Figure 4.6f]. 78

5.1 Model list and the name is abbreviation 105

B.1 Fitted parameters in different LAD model 151

1. INTRODUCTION

The main theme of this research is understanding of key mechanisms of land surface processes and their role in hydrologic cycles and exchange of greenhouse gases, from the soil-vegetation system to land-atmosphere interactions. Here, we explore how complex surface systems are currently conceptualized for modeling and how these conceptualizations perform when tested against field data. Also, we seek improvement of a model for the better understanding land surface system.

Land-surface process plays a vital role as a component of Earth system through the exchange of heat, water vapor, and carbon dioxide from the terrestrial surface to the atmosphere. Substantial and non-stationary change of vegetation/land-use by season and intensity of human activities can significantly influence these processes (Kume et al., 2011; Fisher et al., 2009; Loescher et al., 2005; Sheil, 2018). Including those anthropogenic/seasonal influences, dynamic interactions between the land surface and the atmosphere extensively affect the current state of global atmospheric dynamics. The LSMs have been increasingly studied for future climate predictions but also used to examine climate impacts on the Earth surfaces. For reliable global-scale simulation, LSMs should able to reflect any possible surface environment varied by topographical/regional/seasonal difference, which requires extensive efforts: large amount of information or samples, gathered from different ecosystems, is needed to increase reliability. We need to note that obtaining information about the cause and effect of climate change relies on model simulation, and that it is used for critical purposes such as climate change mitigation policy (Bonan, 2008). Hence, developing accurate LSMs still remains a critical task for Earth system science.

In LSMs, modeling of surface-atmosphere interaction typically starts with incoming solar radiation and precipitation, which are the most important exogenic drivers at the terrestrial surface. The prediction of total absorbed energy from the sun, also known as net radiation R_n, and its partitioning is a key process to understand the surface process and to predict each major fluxes (Wang

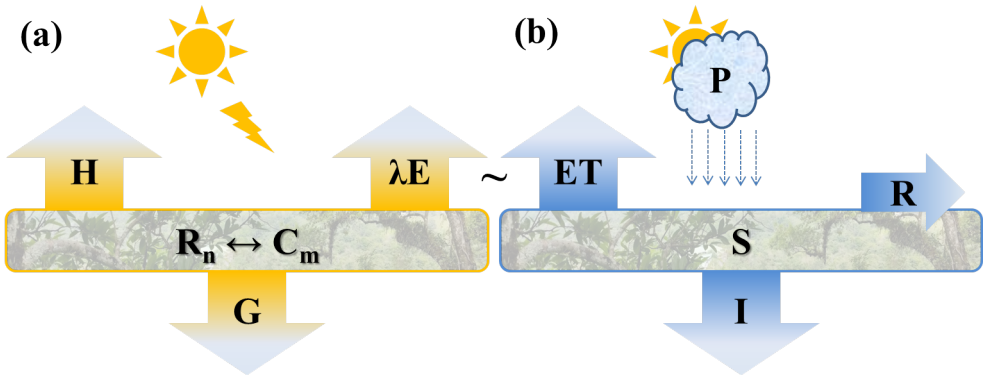

Figure 1.1: (a) Energy Balance, and (b) Water Balance.

et al., 2016) [Figure 1.1a]. The energy balance equation with the main elements is

$$R_n = \lambda E + H + G + C_m \tag{1.1}$$

where λE is latent heat exchange vaporization, H is sensible heat exchange, G is sensible heat exchange to the ground, and C_m is the other miner energy exchange including the metabolism of energy for photosynthesis and storage flux in canopy air space (Heidkamp et al., 2018; Monteith and Unsworth, 2013). This energy budget indicates that the water vapor flux ($\sim \lambda E$) is one of the key components, which highlight the role of water budget is also important in this system [Figure 1.1b]. The water balance equation is

$$P = ET + I + R + S \tag{1.2}$$

where P is precipitation, ET is evapotranspiration, I is infiltration into soil, R is runoff, and S is a storage term (Brutsaert, 2005). Consequently, these two simplified relationships, water and energy balance, become a foundation of LSM, denoting various static or dynamic activities and their interactions.

Water is intertwined with the most of Earth systems, which indicates its importance but also means it makes modeling difficult. The water vapor flux is one of the key components of the energy

exchange [Figure 1.1b] and the most essential source of any organism and the living environment. For instance, vegetation is readily affected by the climate and particularly by water availability. Conversely, it also has an important role in regulating the amount of the water vapor and other greenhouse gases (e.g., carbon dioxide) via photosynthesis. On the other side, humans and trees also pump to the surface from deep ground, which takes a long time to recharge water. Hence, the water moves through various systems, which results in different temporal/spatial characteristics. These each system even a little linked with the water cannot be easily disregarded to understand the land surface system, and consequently the complex characteristics of and interactions between different systems are difficult to capture through simple models. As the concepts included are expanded, LSMs become increasingly complex and must represent multiple constituents (e.g., by matters), spatial scales (e.g., hillslope, watershed, continent), and time periods (e.g., hours, seasons, decades).

Although considering numerous sub-systems is possible for an advanced model, here, for practical purposes, we split the focus into two to three main categories: soil (vadose zone), vegetation, and surface boundary layer. One of the reasons to separate the system is the significantly different time scale (Wang et al., 2006). For instance, precipitation quickly reaches the ground while infiltration is a slow process, thus surface water tends to respond faster to the weather conditions than does deep soil. The vapor state within the surface boundary layer has a faster response to the turbulence transfer than the liquid state in soil or vegetation. Besides, the soil has a complicated system itself to influence a water flux as infiltration or evaporation process, because the water flows through irregular pore space and competing reciprocal actions (Mein and Larson, 1973; Shokri and Salvucci, 2011; Shokri et al., 2008). Vegetation transfers the soil water into the surface through transpiration and carbon uptake for photosynthesis, responding to the intensity of sunlight and atmosphere demands for water. The vegetation characteristic also changes dynamically in spatial and time scale (e.g., tree growth and mortality) (Fisher et al., 2018; Medvigy et al., 2009; Bonan et al., 2014).

Therefore, it is important to keep studying the soil and vegetation processes for the better

understanding and accurate LSM, mainly through model diagnosis and updating. Among these major systems, we limit our research scope and focus on the soil and vegetation system: (1) Soil evaporation without vegetation, and (2) The matter and energy exchange in the vegetated surface. This distinction is necessary because the existence of vegetation makes the fundamentally alters the relative importance of the controlling variables like a switch. It is consequently divided by which surface system either soil (e.g., evaporation) or vegetation (e.g., photosynthesis) is a main controller. Here, atmospheric conditions are important in both cases, but the condition of soil surface (e.g., soil surface temperature, water contents, pore size) can be the main driver for the soil evaporation without vegetation. On the other hand, vegetation forms the canopy which results in shaded soil surface (small energy exchange at the surface), and it conveys soil water through sap-flow suppressing soil evaporation. Moreover, the vegetation is a key contributor to the exchange of other elements (e.g., carbon). With dense canopy, the vegetation would become the main player as the substance and energy exchange, although soil activity still exists.

To investigate those important elements in this study, the first section focuses on the interaction between bare-soil evaporation and sub-surface behavior. Then, the second section of this study addresses canopy processes in a point-scale LSM using Community Land Model (CLM). This study is related to our research project about mountain rainforest in Costa Rica, so it allows us to use extensive observations for the model improvement.

In the first section (Section 2), I focus on how shallow sub-surface conditions interact with evaporation rates, particularly in the very early stages after wetting events. Through this study, a new relationship for the saturated front depth (also known as the characteristic length) is derived under the condition of a shallow water table. Because the front depth in a soil is highly related to the evaporation rate at the surface, the more accurate estimation provides insight into improvement of soil evaporation theory.

Section 3 examines the influence of pore-size heterogeneity on evaporation rates. Here, relative evaporation models are explored, and their improved formulas are introduced to reflect heterogeneous pore sizes at soil surface based on a capillary theory and a pore-size distribution. This model

is tested through a soil-atmosphere resistance scheme (Haghighi et al., 2013), whereby drying rates are controlled by resistances in a soil layer and a viscous boundary layer at the soil-atmosphere interface. Using a heterogeneous representation of wetted soil pores, this exploration give a great insight of its role in between atmosphere and soil system.

In Section 4, a state-of-art LSM is explored, which contains complex physiological and hydrologic processes. Here, the performance of the Community Land Models (CLM4.5 and CLM5) is examined against tower and ground measurements from a tropical montane rainforest in Costa Rica. The study site receives frequent and intense precipitation and experiences high levels of relative humidity. A measurement tower there is equipped with an eddy-covariance and vertical profile systems able to capture various micrometeorological variables, particularly for wet and complex terrain. In this work, results from point-scale simulation for both CLM4.5 and its updated version (CLM5) are compared to observed canopy flux and micro-meteorological data. Through this deep examination, we suggest several improvements to the advanced land surface model.

Finally, in Section 5, more complicated LSM (Multi-Layered CLM, CLM-ml) is tested and modified, to explore how the representation of in-canopy vegetation structure affects model's performance, against the same measurements in Costa Rica along with a vertical canopy profile tower. For further investigation, several canopy shapes are applied via a statistical distribution, and their "true" shape (more complex shape) is fitted using mixed-distributions. To reflect the canopy shapes in CLM-ml, a first-order-closure model is applied for turbulence. This study highlights the importance of in-canopy characteristics for the accurate prediction of forest-atmosphere interactions.

2. SATURATED DEPTH (CHARACTERISTIC LENGTH) OF POROUS MEDIA UNDER STEADY-STATE EVAPORATION

2.1 Introduction

Evaporation is one of the key processes in the hydrological cycle, with obvious linkages to other biogeochemical cycles. The interplay between evaporation and precipitation determines the amount of water available for other hydrologic processes (such as runoff or subsurface transport processes). Because the evaporative flux also acts as an important lower boundary condition for global meteorological simulations, the evaporation rate is a vital component for understanding current and future climate. In contrast to the evaporation from a free-water surface (such as a lake), the physics of evaporation from porous media has complicated characteristics which are incompletely understood. Conversely, evaporation from bare soil avoids the complications present in transpiration where individual plant species responses biologically control the water flux rate. Natural soils have irregular pore sizes, and significant spatial heterogeneity, so a-priori prediction of evaporation from bare soil systems remains a challenge, and often depends on empirical calibrations. Because the evaporation rate of a soil system depends on both the soil characteristics as well as the atmospheric demands for water, The dynamic interactions between the supply of water and atmospheric demand result in an extensive variety of evaporation behaviors through time (Lehmann, 2008).

During the evaporation process at the soil surface, the relative evaporation rate stays in nearly constant while surface water decreases, and such process is not fully understood yet due to its complexity. This relative drying rate tends to show different stages which correspond to periods where different processes are in control of the evaporation rate. At the beginning of evaporation with a fully saturated condition, the dying rate remains constant at the potential drying rate (Stage I) until the condition reaches 'critical water content' at the surface (Sherwood, 1930). After that, its rate starts to decrease; this period is called as Stage II. Based on this idea, two or three steps of evaporation type are defined (Idso, 1974; Philip, 1957).

Many studies parameterize such this Stage I/Stage II phenomenon using empirical methods (Kondo et al., 1990; Chanzy and Bruckler, 1993; Griend and Owe, 1994; Swenson and Lawrence, 2014). However, an alternate approach has been introduced by some studies to more physically interpret this phenomenon (Suzuki and Maeda, 1968; Schlünder, 1988), in which the water vapor flux from the soil surface is controlled by the wind layer's thickness (viscous boundary layer, VBL) and pore size. This conceptualized model based on diffusion theory can represent Stage I and Stage II together, and it proves the existence of a constant drying period.

On the other hand, to predict evaporation, an understanding of the mechanisms taking place inside soil system (below VBL) is also vital, because the source of the evaporation comes from inside of the soil. The water movement inside a soil system is governed by a number of possible potentials, but the most important drivers are known to be the gravitational potential and matric potential (which is also called capillary potential) (Hillel, 2003). The soil drying rate is highly related to the two important activities depending on which potential dominates the total potential. The negative pressure (suction) of a partially saturated soil leads to redistribution of moisture in the soil, which is a key mechanism for continuously furnishing a top soil layer with the evaporative water. Due to the physical relationship between the pore size distribution and the negative pressure (Hillel, 2003), some models for the soil water have been derived based on capillary theory and pore size distribution; one such representative example is the unsaturated hydraulic conductivity model (Mualem, 1976).

Hence, many studies have emphasized the importance of soil characteristics to understand and mimic the evaporation mechanism (Shokri et al., 2008; Shahraeeni et al., 2012; Haghighi et al., 2013). Haghighi et al. (2013) showed the evaporation rate is controlled by the viscous boundary layer (VBL) but that the water supply is also limited by soil resistance - conditions both below and above the soil need to be taken into account. Using resistance terms based on the Schlünder model and Darcy's law, Haghighi's results compared well to experimental data, and reduced the discrepancy of the diffusion-only model for VBL. In this case, the evaporation rate is related to atmospheric demands, which are dominated by diffusion process (Haghighi et al., 2013), and the

water for the evaporation from the spil surface is supplied from the soil, and hence is controlled by soil properties.

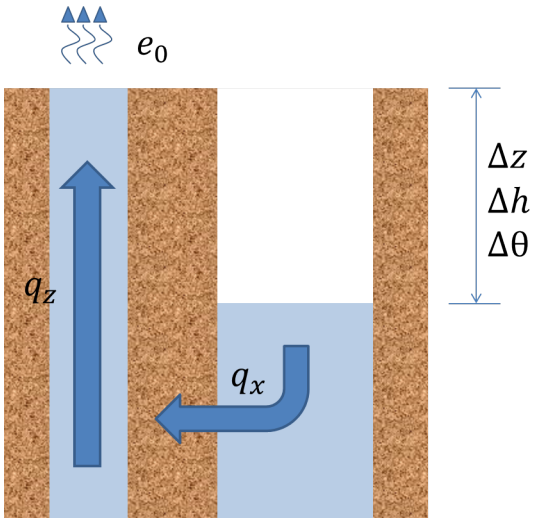

Figure 2.1: Water supply machanism for drying soil in pore scale. Δz, Δh, and $\Delta\theta$ are the vertical difference of height, pressure head, and the water contents. q_z, q_x, and e_0 are vertical water flux, horizontal water flux, and surface evaporation flux. The Figure is modified from (Lehmann, 2008)
.

Particularly, to track such the constant soil evaporation (Stage I), the soil characteristics for water flows are an important factor. Yiotis et al. (2006) investigated the constant drying rate (Stage I) using numerical simulations, and found that pore size and the pore size distribution are related to the constant drying rate. Higher Bond number systems, where the gravitational force is larger than surface tension, tend to have shorter periods of Stage I (Yiotis et al., 2006). Lehmann (2008) also studied the Stage I evaporation, and used the relation between the duration of Stage I and the evaporative front depth, which is also known as the characteristic length. This study expressed the front depth as a function of capillarity, gravity, and viscous forces. The derivation of this relationship is based on capillary potential or matric potential having a role of supplying the water onto a surface (Lehmann, 2008). Here, the assumption was that the soil-water system consists of thin films. The use of only capillary potential cannot be realistic since the matric potential is

the bigger concept for soil system (Hillel, 2003). However, their analysis and results show that the behavior of soil water based on capillary theory using the scheme of two films [Figure 2.1] are similar to those based on matric potential, which indicates pore size and its distribution are important to determine the front depth. In Lehmann (2008), the derivation of the characteristic length is simplified for practical calculation and it is simulated well for relatively large particles such as sand.

The saturated depth (front depth or characteristic length) essentially implies the characteristics of surface water content and soil, which are closely related to evaporation rates. Recently, many studies have investigated the use of the characteristic length concept in the development of models for the movement of vapor and liquid water (see Shahraeeni et al. (2012); Sadeghi et al. (2012); Neriah (2014); Sadeghi et al. (2014)), because the characteristic length enables the estimation of evaporation using their correlations. For examples, applying the characteristic length theory, derived from assumptions about vertical diffusion coming from dried pores below the soil surface, and the Suzuki and Maeda (1968) model, which addresses vertical and horizontal direction transfers within a boundary layer above the soil surface, time series of evaporation rates were well modeled by Shahraeeni et al. (2012). Their results show the front depth has an important/useful role in determining the water supply at the soil surface. Relationships such as this can also be used in global simulations using Community Land Model (CLM5). The term used in the CLM work is different (Dry Surface Layer - DSL), but the concept is the similar and it gives great improvement on soil evaporation (Swenson and Lawrence, 2014; Lawrence et al., 2018). Sadeghi et al. (2012) developed an exact solution and a simplified solution to the front depth (Sadeghi et al., 2012), applying an assumption of the shallow water table and considering the vertical transport of water supply based on Darcy's law and Brooks-Corely unsaturated hydraulic conductivity (Brooks and Corey, 1964).

The front depth has been taken to be formed by the single direction flow like the usual 1D simulation of soil moisture transport. However, it can be argued that the characteristic length is more likely controlled by a two dimensional flows. This is mainly because the retention curve seems

to provide volume changes (via volumetric water contents $\propto$ pressure head) within Darcy's law during the evaporation process, but the Darcy's flow focuses on a pressure head difference. The volume estimated from the head still represents the pressure not actual volume yet. Therefore, the Darcy's law without additional constraint cannot promise the mass conservation. Considering only 1D flow may not be adequate to represent the realistic behavior of the characteristic length. Based on an idea of Lehmann (2008), if a soil system consists of films with various pore sizes, the water movements are mainly governed by capillary force. In the film zone, there can exist both thick and thin films of water [Figure 2.1], which tend to have different pressure heads. Then, the water held by thin films would reach the surface to supply the surface evaporation. At the same time, the water in the thin films lost through the surface evaporation is refilled from the adjacent thicker films due to the different pressure heads. This water re-supply is viewed to take place in a horizontal direction (contrasting with the vertical flow to the surface), and an equation for water movement in the horizontal direction needs to be included (Lehmann, 2008). In Lehmann (2008), for the derivation of characteristic length an empirical equation, the van Genuchten-Mualem (VGM) model (van Genuchten, 1980; Mualem, 1976), was used as a retention curve. On the other hand, the derivation of Sadeghi et al. (2012) is based solely on vertical direction flow, assuming the one directional soil system. Although Sadeghi et al. (2012) used the different retention curve and hydraulic conductivity equations (Brooks and Corey (BC) equation (Brooks and Corey, 1964) compared to Lehmann (2008), their mathematical approach to the derivation of the characteristic length is identical. However, it is important to keep in mind their intrinsic ideas differ on whether evaporative water is supplied vertically or horizontally.

Normally, the front depth is directly estimated using Darcy's law without applying the conservation of mass to avoid having to solve Richards equation numerically (Richards, 1931). Nevertheless, the question here is whether the front depth estimated using a single directional Darcy's equation can accurately represent accurate saturated depths. In this study, a comparison is made whether the estimated front depth is different if conservation of mass is also used. To lead to better agreement for the simplified model, the front depth theory is updated based on the approaches of

Lehmann (2008) and Sadeghi et al. (2012) under the steady state of evaporation (Stage I). Simulations may show the suggested method would not only work well for the soil having somewhat uniformly-distributed-particle sizes (large n for VGM), but it is also suitable for the soils having inconsistently-distributed-particle sizes (small n). Comparisons will also be conducted between a previous method (Lehmann, 2008; Sadeghi et al., 2012), the new method, and a numerical simulation of Richards equation. We also discuss some of the difficulty in simulating the front depth when the parameter 'n' is small, and point out how the new approach avoids this difficulty.

2.2 Methodologies

2.2.1 Theory

In this section we derive a new 2-dimensional flow approach for the estimation of characteristic length. The derivation considers one grid box with large-scale flow occurring in a single dimension (z-axis) flow q_z on the larger scale. The assumption is that there is no horizontal flow between grid boxes $q_x = q_y = 0$. Therefore, any horizontal flow shown in here is fine scale flow within a single grid box. The scale we consider here would be smaller than Darcy REV-scale which is normally used for field studies, but still larger than pore-scale because it is based on Darcy's law. Then, with conditions that soil water flows out of the soil system only in vertical direction, an anisotropic Darcy's law can be written as

$$q_z = -K_{xz}(h)\frac{\partial h}{\partial x} - K_{yz}(h)\frac{\partial h}{\partial y} - K_{zz}(h)\frac{\partial h}{\partial z} = 0 \tag{2.1}$$

$$q_z = q_{xz} + q_{yz} + q_{zz} \tag{2.2}$$

where q_z is total vertical flow, x and y are the horizontal directions, z is the vertical direction, K is the unsaturated hydraulic conductivity (K_{xz} indicates flow direction from x to z), and h is pressure head. In Eq. (2.2), $q_{yz} + q_{zz}$ indicates horizontal flow joining vertical flow, and q_{zz} is vertical flow. If there is no any interference from horizontal flow ($q_{xz} = q_{yz} = 0$), q_z becomes equal to q_{zz}. It

means there is no horizontal flow both within and between a grid box.

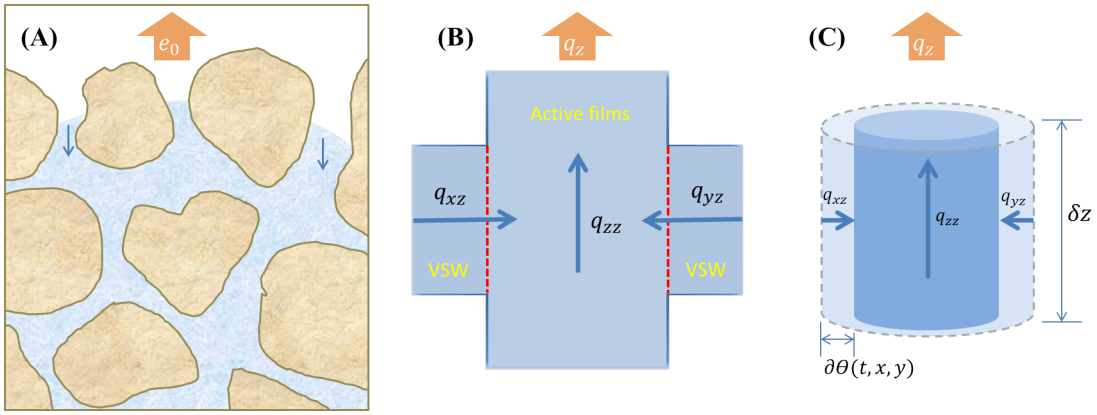

Figure 2.2: The scheme of evaporation process. With the idea that the evaporation rate occurs only at active pores (thin films), the water under the inactive pores which is called as vertically stable water (VSW) in this manuscript can migrate through vertical flux under the active pores.

The underlying idea used in past derivations of characteristic length is that the supply of the water moves vertically to the soil surface due to difference of capillary head or matric head as Sadeghi et al. (2012). Based on Darcy's law, vertical flow can be written as

$$q_z = -K_{zz}(h)\left(\frac{\partial h}{\partial z} + 1\right) \tag{2.3}$$

where K_{zz} is the unsaturated hydraulic conductivity in the of vertical direction. This equation has the same form to the anisotropic equation with $q_{xz} = q_{yz} = 0$. In this condition, dotted lines in [Figure 2.2b] for horizontal flow are closed. Here, h can be directly related to water volume using an empirical relationship (e.g., retention curves). However the head in the equation represents the pressure, so the equation cannot necessarily insure mass-conservation.

To properly estimate the front depth and hold mass-conservation, we want to consider the case where the horizontal flow within the grid box may not be zero $q_{xz} = q_{yz} \neq 0$. In other words, if horizontal flows were allowed, the dotted lines in [Figure 2.2b] become open, and q_{xz} and q_{yz} are no longer zero. At this time, the amount of horizontal flux $q_{xz} + q_{yz}$ is unknown but we assume that

its flow path is known. In this system, all water continues to eventually exit through one vertical pipe (small pore). Therefore, it should hold $q_{xz} + q_{yz} = q_{zz}$ as an equilibrium, as long as $q_{xz} + q_{yz}$ allows the evaporation rate e_0 or q_{zz}. This also supports the idea that water can be taken to move in the horizontal direction from a large pore into a nearby small pore, due to the flow which is forced by their vertically different matric forces (Lehmann, 2008). In this schema the horizontal flows should not contribute to long-distance water movement since the flows occur between adjacent soil pores within the grid box. Smaller pores lose the water via evaporation (vertical direction) and this active flow produces a vertical pressure difference. At the same time, the water recharges from the surrounding bigger pores which have self immovable status by the vertical suction force. We will call here this status as vertically stable water (VSW), and this water mainly flows out through a smaller pore as long as the water is still connected in the horizontal direction [Figure 2.2b]. Therefore, these horizontal flows can be triggered by vertical flow which caused by a difference between two heads which is equivalent to Δh in the vertical direction (Lehmann, 2008). The absolute vertical length (z) of the water migration from a large pore to a small pore is equal to Δz [Figure 2.1; Figure 2.2], and so this water migration through a horizontal way is equivalent to vertical flow by $\Delta h/\Delta z$, so hydraulic conductivity would also follow $K_{zz}(h)$ if the system is isotropic. The combined effects of this process may form the front depth in soil system. Finally, vertical water flow exists coming from lower grid in 1D system too. Therefore, the total resistance by a soil system can be lower than our previous thought.

The equilibrium $q_{xz} + q_{yz} = q_{zz}$ can be also derived from Richards equation.

$$\frac{\partial \Theta}{\partial t} = -\frac{\partial q_z}{\partial t} = \frac{\partial}{\partial z}\left[K_{zz}(h)\left(\frac{\partial h}{\partial z} + 1\right)\right] \tag{2.4}$$

If the water contents in the vertical direction were always wet (the water film is not vertically disconnected) and the total water contents only changed in horizontal dimensions [Figure 2.2c], z can be independent from other dimensions in $\Theta(t, x, y, z)$. Taking a vertical integration (z direction)

in the both sides with small vertical distance δz, the Richards equation can be written as

$$\frac{\partial \Theta(t,x,y)}{\partial t}\delta z \approx \int_z \frac{\partial \Theta(t,x,y,z)}{\partial t} dz = K_{zz}(h)\left(\frac{\partial h}{\partial z}+1\right) \tag{2.5}$$

Finally, as in the left hand side, the change of water content $\Theta(t,x,y)$ with respect to time, within a certain time and certain vertical length δz, becomes only a function of the horizontal direction. According to this relationship, the change of the water amount (in length scale here) is directly related to the change of horizontal size of water content $\partial\Theta(t,x,y)/\partial t$, and it is equal to the vertical flux as in the right hand side [Figure 2.2c]. Therefore, we can substitute the horizontal flux term in the right hand side as

$$K_{xz}(h)\frac{\partial h}{\partial x} + K_{yz}(h)\frac{\partial h}{\partial y} \approx K_{zz}(h)\left(\frac{\partial h}{\partial z}+1\right) \tag{2.6}$$

Again, this shows the equilibrium $q_{xz} + q_{yz} = q_{zz}$ is necessary for the mass conservation. The left term represents horizontal fluxes which finally flow out via vertical flow (evaporation) and it also represents the mass change and size change of water contents. We also need to note that the left hand side equation is derived from $\partial\Theta/\partial t$, so the flow occurs in one grid box.

If the horizontal flow or the change of the water volume $\partial\Theta/\partial t$ is only transported through the same path as the vertical flux q_{zz}, Eq. (2.2) to conserve the mass becomes

$$q_z = -c \cdot K_{zz}(h)\left(\frac{\partial h}{\partial z}+1\right) \approx -2 \cdot K_{zz}(h)\left(\frac{\partial h}{\partial z}+1\right) \tag{2.7}$$

where c is a constant for when the two relationships (two direction flows) are not identical in Eq. (2.6), and the c may become 2 if they are identical. However, the exact value of c is unknown yet. The important thing to note here is that the horizontal flux in Eq. (2.6) cannot exceed the vertical flux because the horizontal flux depends on the vertical flux and the vertical flux is determined by the evaporation rate. Here, $c \cdot K_{zz}(h)$ does not mean the over flow of q_z, but it rather implies lowered resistance against q_z. If this is the case, the parameter c could be close to 2. On the other

hand, the parameter c could be less than 2, if the horizontal flux is limited by a soil property such as very low horizontal saturated K_s value. Also, c becomes 1 (considering only a single direction of flow like as Eq. (2.3)), if saturated K_s is zero

2.2.2 General Formulation

The front depth can be estimated by solving Eq. (2.7). If the total flux q_z is given by the evaporation rate e_0, the front depth using pressure head terms and water contents terms becomes

$$\begin{cases} Z(h_{surf}) & = \int_0^{h_{surf}} 1/\left(\frac{e_0}{c \cdot K(h)} + 1\right) dh \\ Z(\Theta_{surf}) & = \int_1^{\Theta_{surf}} 1/\left(\frac{e_0}{c \cdot K(\Theta)} + 1\right) \frac{\partial h}{\partial \Theta} d\Theta \end{cases} \tag{2.8}$$

$$\Theta = \frac{\theta(h) - \theta_r}{\phi - \theta_r} \tag{2.9}$$

where c is a constant (1 is single flow and 2 is double flow in this study), Z is the front depth also known as the characteristic length (L_c), Θ is a normalized water content, θ is a volumetric water content, θ_r is a residual water content, ϕ is a saturated water content or a porosity, $\partial h/\partial \Theta$ can be estimated through any retention curve, h_{surf} and Θ_{surf} are the matric potential head and the normalized volumetric water content at the surface (Lehmann, 2008; Sadeghi et al., 2012). Using the normalized water content term is useful because it has a bounded range from 0 to 1 for the numerical integration if it is necessary. The integration would be solved exactly or numerically depending on the unsaturated hydraulic conductivity model $K(\Theta)$. The solution from numerical integration is called here a semi-analytical method because it is much simpler than solving Richards equation.

To account for the total mass loss $V[L]$ to reach Θ_{surf} from the fully saturated condition, Eq. (2.8) adds the water content from the integration to become

$$V(\Theta_{surf}) = \int_1^{\Theta_{surf}} (1 - \Theta) \,/ \left(\frac{e_0}{c \cdot K(\Theta)} + 1\right) \frac{\partial h}{\partial \Theta} d\Theta \tag{2.10}$$

and its cumulative time T from Eq. (2.10) is

$$T(\Theta_{surf}) = V(\Theta_{surf})/e_0 \tag{2.11}$$

Calculation of the total mass lost V is useful for comparison of the model results with numerical simulations of Richards Equation, since it too is derived from mass conservation. Also, the cumulative time which indicates the end of Stage I can be predicted using Eq. (2.11). Sometimes this cumulative time is useful because for some other models it can define the ending point of Stage I evaporation. Comparing between the previous Eq. (2.5) method ($c = 1$) and the new method ($c = 2$) given in Eq. (2.8), the suggested method has only a higher hydraulic conductivity K than the previous method, without any complicated modifications.

2.2.3 Special Case (Closed-Form)

This section provides a closed-form equation using a relatively simple hydraulic conductivity model based on the derived equation in the previous section. Considering Brooks and Corey hydraulic conductivity as (Sadeghi et al., 2012),

$$K(h) = \begin{cases} K_s & (h \le h_b) \\ K_s \left(h/h_b\right)^{-P} & (h > h_b) \end{cases} \tag{2.12}$$

The front depth can be calculated as

$$Z(h_{surf}) = \begin{cases} \frac{h_{surf}}{\frac{e_0}{c \cdot K_s}+1} & (h_{surf} \le h_b) \\ \frac{h_b}{\frac{e_0}{c \cdot K_s}+1} + \int_{h_b}^{h_{surf}} 1/\left(\frac{e_0}{c \cdot K(h)}+1\right) dh & (h_{surf} > h_b) \end{cases} \tag{2.13}$$

where h_b is the head of air entry pressure, P is a Brooks and Corey parameter, c is 2 for our case and 1 for previous case, and K_s is saturated hydraulic conductivity. The solution for the integration part in Eq. (2.13) follows the method described in Sadeghi et al. (2012), and the derivation is based on assumption of $e_0 < c \cdot K_s$ or $e_0 < c \cdot K(h)$ (Sadeghi et al., 2012). Using a Maclaurin series

expansion, the integration can be

$$\int_{h_b}^{h_{surf}} 1/\left[-\left(-\frac{e_0}{c\cdot K(h)}\right)+1\right]dh = \int_{h_b}^{h_{surf}} \sum_{n=0}^{\infty}\left[-\frac{e_0}{c\cdot K_s}\left(\frac{h}{h_b}\right)^P\right]^n dh = \sum_{n=0}^{\infty}\frac{h_{surf}}{Pn+1}\left[-\frac{e_0}{c\cdot K_s}\left(\frac{h_{surf}}{h_b}\right)^P\right]^n - \sum_{n=0}^{\infty}\frac{h_b}{Pn+1}\left[-\frac{e_0}{c\cdot K_s}\cancel{\left(\frac{h_b}{h_b}\right)^P}\right]^n \tag{2.14}$$

The infinite series is equivalent to the Gaussian hypergeometric function, allowing Eq. (2.14) to be rewritten as

$$\int_{h_b}^{h_{surf}} 1/\left[-\left(-\frac{e_0}{c\cdot K(h)}\right)+1\right]dh = h_{surf}\cdot 2F1\left(1,\frac{1}{P},1+\frac{1}{P},-\frac{e_0}{c\cdot K_s}\left(\frac{h_{surf}}{h_b}\right)^P\right) - h_b\cdot 2F1\left(1,\frac{1}{P},1+\frac{1}{P},-\frac{e_0}{c\cdot K_s}\right) \tag{2.15}$$

where $2F1$ is the Gaussian hypergeometric function. The definition of the hypergeometric function and detailed steps of conversion, from the infinite series like as Eq. (2.14) into $2F1$, are described in the Appendix A. In general, the coefficients of the resulting hypergeometric function do not allow to be reduced a specific special function (i.e. Bessel function), but numerical evaluation of the Gaussian hypergeometric function is straightforward in modern computer programming systems such as Matlab or Python. Based on Eq. (2.15), some simulation results are also presented in [Figure 2.3].

Additionally, the total volume of mass loss can be estimated as

$$V(h_{surf}) = \begin{cases} 0 & (h_{surf}\le h_b) \\ \int_{h_b}^{h_{surf}} (1-\Theta(h))/\left(\frac{e_0}{c\cdot K(h)}+1\right)dh & (h_{surf}>h_b) \end{cases} \tag{2.16}$$

where $\Theta(h)$ is $(h/h_b)^{(2-P)/3}$ here and the integration can be estimated in terms of the hypergeo-

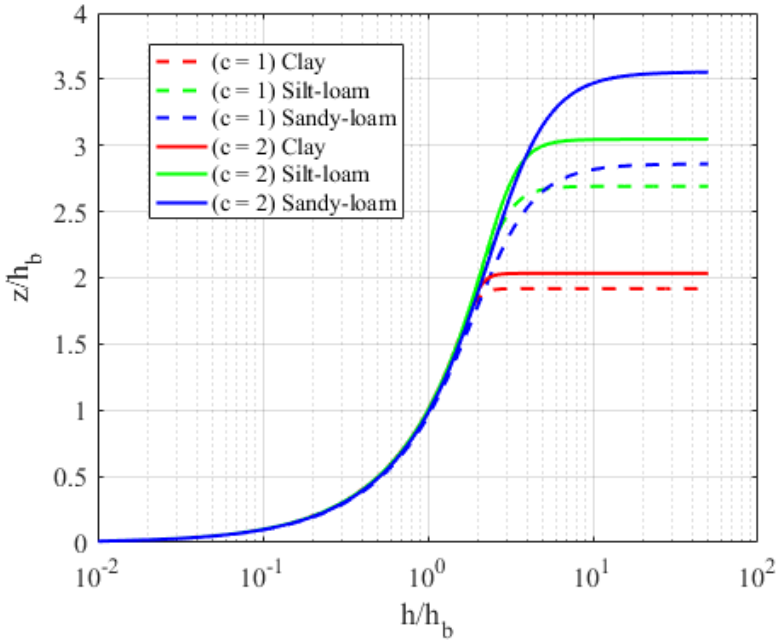

Figure 2.3: Test Simulation using Eq. (15): $c = 1$ (Sadeghi et al. (2012)) and $c = 2$ (New method).

metric function as

$$\begin{aligned}\int_{h_b}^{h_{surf}} (1-\Theta(h))/\left(\frac{e_0}{c\cdot K(h)}+1\right) dh \\ = \left[h\cdot\left(2F1\left(1,\frac{1}{P},1+\frac{1}{P},-\frac{e_0}{c\cdot K_s}\left(\frac{h}{h_b}\right)^P\right)\right.\right. \\ \left.\left.+\frac{3}{P-5}\left(\frac{h}{h_b}\right)^{\frac{2-P}{3}} 2F1\left(1,\frac{P-5}{3P},\frac{2P+5}{3P},-\frac{e_0}{c\cdot K_s}\left(\frac{h}{h_b}\right)^P\right)\right)\right]_{h_b}^{h_{sulf}}\end{aligned} \tag{2.17}$$

To estimate the maximum depth $Z_{max}(= L_c)$ (the depth when hydraulic disconnection occurs) for this new model in simplified version, we can follow the derivation steps outlined in Sadeghi et al. (2012), and find:

$$\begin{aligned}Z_{max} =& h_b/\left(\frac{e_0}{c\cdot K_s}+1\right)+h_e-\frac{h_e}{P+1}ln(2)-h_b \\ &+\frac{h_b}{P+1}ln\left(\frac{e_0}{c\cdot K_s}+1\right)-h_e\left(\frac{ln(2)}{1-P}+\frac{\pi^2-12ln(2)}{12P(1-P)}\right)\end{aligned} \tag{2.18}$$

where he h_e is equal to $h_b \left[e_0/(c \cdot K_s)\right]^{-1/p}$. This expression is an approximate solution but it provides reasonable results (Sadeghi et al., 2012).

2.3 Simulation Setup

First, while ideally comparison of the models would be made against field measurements, the fine scale of vertical resolution of soil moisture needed for comparison with the model is not readily attained. To allow, therefore, between a single-flow (old method) and double-flow (new method) approaches for the front depth Eq. (2.8) are first compared using the Richards equation Eq. (2.4) as "true" data set. The Richards equation requires numerical method to solve, so it is accomplished by the aid of a HYDRUS 1D program for this test. HYDRUS 1D solves the uses a finite element approach for the numerical solution of Richards equation in one dimension. The Richards equation is derived based on mass-conservation, so it can be a good target for this comparison. Since HYDRUS 1D does not use the simplifications of the Richards equation, it holds the potential to match better the true physical conditions of the soil moisture flow. The simulations for this study were designed to estimate the characteristic lengths and to enable comparison between the analytical solutions and the numerical solution. For that, both the HYDRUS 1D program and the two characteristic length solutions use the same retention curve equation, which requires soil physical parameters. Given the choice of the Van Genuchten-Mualem model (VGM) for the hydraulic conductivity equation, parameters n, α, and K_s are required. Therefore, following methods for those parameters are taken from Carsel and Parrish (1988): (the r^2 for these coefficient values are 0.94 and 0.92 (Lehmann, 2008).

$$\alpha = 0.0087 \cdot (n - 1) \qquad [mm^{-1}] \tag{2.19}$$

$$K_s = 0.0077 n^{7.35} \cdot 10^3 \qquad [mm \cdot day^{-1}] \tag{2.20}$$

where n is a VGM parameter related to a pore size distribution, α is the inverse air entry pressure h_g, K_s is the saturated hydraulic conductivity.

For the HYDRUS 1D simulation as a "true" result, the setting of the simulation mimics the way Lehmann (2008) did. An initial condition of $h = -1mm$ was set for the entire simulation domain; this is equivalent to saturation. The simulation domain was set at a depth of $3m$, and 501 nodes were used in the calculations. For the lower boundary, a zero-flux boundary condition was used. For the upper-boundary flux, again following the simulations of Lehmann (2008), a constant evaporation rate of either $3mm/day$ or $10mm/day$ was set to represent low and high atmospheric demands. Using the simulation results from HYDRUS 1D, additional analysis is made to estimate the front depth at the end of Stage I. For numerical reasons which will be discussed below, the drying front depth in the simulations was defined as the depth where the volumetric water content was equal to 0.99 of saturation. Additionally, because the condition of the end of Stage I should be defined, the criteria was set as 15 bars at the soil surface (Lehmann, 2008; Kozak et al., 2005).

For soil-water retention curve and hydraulic conductivity models, the van Genuchten-Mualem models can be written as

$$\Theta = \left[1 + (\alpha h)^n\right]^{-(1-1/n)} \tag{2.21}$$

$$K(\Theta) = K_s \Theta^\tau \left[1 - \left(1 - \Theta^{\frac{1}{1-1/n}}\right)^{1-1/n}\right]^2 \tag{2.22}$$

where n is a parameter related to a pore size distribution, $\aleph$ is equal to $1/h_g$, K is the hydraulic conductivity, K_s is the saturated hydraulic conductivity, and τ is 0.5. Unlike Sadeghi et al. (2012), who used the Brooks and Corey equation for the soil retention curve and the Gardner equation for the unsaturated hydraulic conductivity, the integral forms of Eq. (2.8),and Eq. (2.10) based on VGM models cannot be exactly solved, and so numerical integrations are employed for this simulation. However, this loss of an analytic solution is outweighed by the value of the greater verification and acceptance of the VGM models. Moreover, Eq. (2.8) and Eq. (2.13) are not significantly sensitive to the method use to conduct the numerical integration for this simulation. The shapes of the functions are not complicated (normal or exponential shape) and, additionally, the integral range is small if Θ domain is considered. This method provides a much lower com-

putational time compared to a complete numerical solution of Richard's equation. For example, to solve Eq. (2.8) and Eq. (2.13), 20 nodes in one time were enough to calculate the integration but about 500 nodes in many times (iterated through time) were necessary in HYDRUS for this study. It should be noted that because, maximum front depths and time are usually unknown, a trial-and-error approach is often needed to find the proper number of nodes and maximum depths and time for the HYDRUS model.

For the comparisons, various plots will show simulation results of the front depth and total volume/time at the end of Stage I, with different soil properties 'n'. The first simulation is conducted based on a single direction flow Eq. (2.8) with ($c = 1$) to test the previous theory and the results are compared with the results of Richards equation Eq. (2.4) under HYDRUS 1D. Due to that the Richards equation is based on mass conservation, the cumulative times from Eq. (2.10) are also compared. Then, the new methods given in Eq. (2.8) and Eq. (2.10) with ($c = 2$) are simulated and the results are compared in the same way.

Additionally, the experimentally-based data from the literature (Sadeghi et al., 2012; Shokri and Salvucci, 2011) are compared using the exact solution which is derived based on Brooks and Corey hydraulic conductivity in the section 2.1. Except newly introduced parameter ($c = 2$), the comparison way and calculation methods (other soil parameters) are identical with the previous study (Sadeghi et al., 2012).

As long as concerning about the terminology and abbreviation, we use $L_c = Z_{max}$ as the saturated depth at the end of Stage I. Also, Z should be between 0 and Z_{max} based on Eq. (2.8). The characteristic length (L_c) is used in (Lehmann, 2008), and it is also called as the maximum water table depth (D_{max}) in (Sadeghi et al., 2012). These all represent a depth from a surface to the fully saturated location in the soil system. Therefore, it can be also possibly called as front depth, wet front, saturated depth, etc.

2.4 Results and Discussion

In the comparison, the previous method cannot adequately predict the front depth, although it tends to follow the behavior (HYDRUS 1D) driven by different soil properties [Figure 2.4]. Using

the front depth equation based on single-direction flow Eq. (2.8) with ($c = 1$) results in low front depths compared to the Richards Equation Eq. (2.4) simulated by HYDRUS 1D. As was found in the previous study (Lehmann, 2008), the characteristic lengths (L_c) are well simulated if the soil parameter 'n' is large ($n > 2$). Overall, the trend between n and maximum drying front depth (L_c) is well predicted, although there is again significant divergence of the predictions at small 'n' values. Like the front depth, the total duration until the end of Stage I, which is related total mass loss Eq. (2.10), are not well predicted [Figure 2.5].

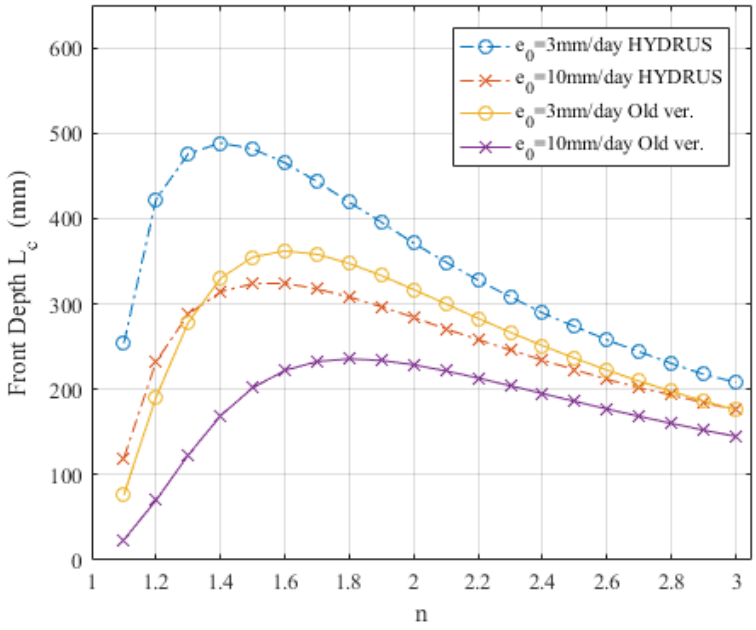

Figure 2.4: Estimated front depth ($L_c = Z_{max}$) vs VGM parameter 'n' where e_0 is 3 mm/day and 10 mm/day. Dotted lines are estimated by HYDRUS and general lines are calculated based on using previous method Eq. (2.8) with ($c = 1$). The Root Mean Squared Errors (RMSE) of the prediction in this graph are 12.114 mm for 3 mm/day and 3.151 mm for 10 mm/day.

On the other hand, the suggested two-direction flow method Eq. (2.8) with ($c = 2$) introduced in this study shows better agreement [Figure 2.6] for the front depth, compared with [Figure 2.4]: the Root Mean Squared Errors (RMSE) are decreased by 85% and 68% in each $3mm/day$ and $10mm/day$ simulation. In the results of total duration and cumulative mass losses, simulations

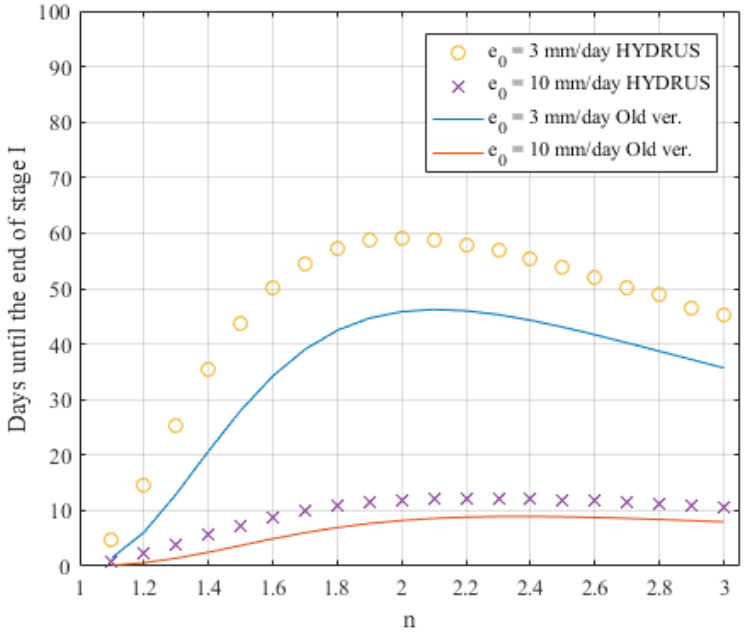

Figure 2.5: Days for the end of Stage I vs VGM parameter 'n'. Lines are earned through Eq. (2.10) and Eq. (2.11) with ($c = 1$) and marks from HYDRUS.

estimated by the newly presented method agree well with those given by the Richard's equation simulation [Figure 2.7;Figure 2.8]. These tests indicate that the new theory gives an improvement of the estimation of the front depth and the conservation of mass. This result reveals that the formation of the saturated depth are not mainly controlled by only single direction flow, but they are rather affected by both horizontal redistribution and vertical direction flows.

Newly introduced method can reasonably predict the front depth as Z_{max} or L_c at the end of Stage I, but it also shows possible predictions according to the state of surface water contents as $Z(\Theta_{surf})$. An additional example supports the idea between surface water contents and the front depth using the semi-analytic solution [Figure 2.9]. It shows, although the results are not perfectly matched, the method introduced in this study has ability to predict the front depth depending on the surface water content.

An interesting aspect in [Figure 2.9] is that soils with very small 'n' (low homogeneity) tends to have bigger discrepancy in the comparison, although the front depth at the end of Stage I as in

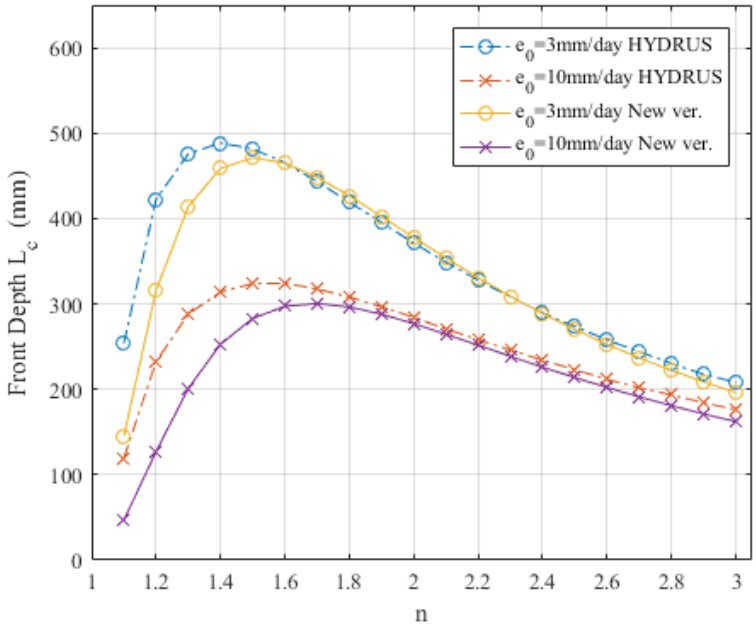

Figure 2.6: Estimated front depth ($L_c = Z_{max}$) vs VGM parameter 'n' where e_0 is 3 mm/day and 10 mm/day. Dotted lines are estimated by HYDRUS and general lines are calculated based on using previous method Eq. (2.8) with ($c = 2$). The RMSE in this graph are 1.875 mm for 3 mm/day and 0.993 mm for 10 mm/day.

[Figure 2.6] is well simulated with most of soil properties. The errors with small 'n' may result from numerical error, inexact c parameter, or more possibly inadequate definition of front depth ($\Theta = 1$). Also, one might cast doubt to the criteria of the pressure head for the end of Stage I ($-1500kP$) for all soil types.

During this study, precision and accuracy problems arise in numerical method (HYDRUS) to identify the true front depth, which can cause in the bigger error with small n. If we look the profile of depth of the HYDRUS simulation in [Figure 2.10], the boundary of saturated zone tends to reach a deeper depth quickly with a small value of 'n', which makes it difficult to define where the front depth ($\Theta = 1$) is. By increasing the precision in HYDRUS, the deep depth (long tail) easily goes to infinity. Therefore, it may be dubious to define such deep-depth as the front depth or evaporative depth, because this deeper depth does not seem the main contributor of the evaporation (see $n = 1.1 \& e_0 = 3mm/day$ in [Figure 2.10]). As can be seen in in the figure, distinguishing

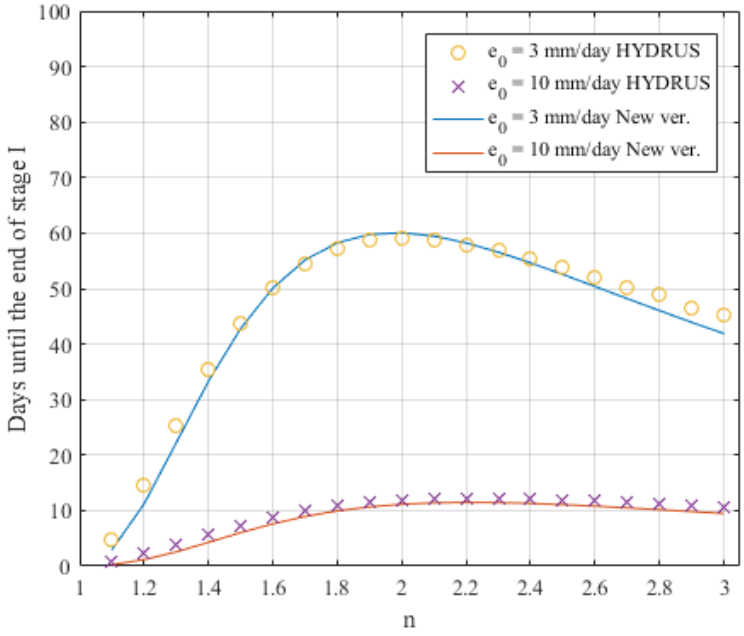

Figure 2.7: Days for the end of Stage I vs VGM parameter 'n'. Lines are earned through Eq. (2.10) and Eq. (2.11) with ($c = 2$) and marks from HYDRUS.

between the depths of $-0.3m$ and $-0.5m$ seems unimportant for determining an exact saturated depth. It is more essential to find the actual contributing depth for the evaporation.

Hence, some criteria may be necessary to use a numerical way. In order to avoid selecting an incorrect front depth caused by the long tail and wasting time for changing precisions, more discretization is usually required to reduce numerical error or to get a more precise solution, which results in a longer computational time. In [Figure 2.6], the front depth is defined at the water content equal to 0.99 instead of 1 for both HYDRUS and methods from Eq. (2.8). On the other hand, a simulation in [Figure 2.11] shows that the front depth is defined at the water content equal to 0.95, which does not show a significant change of the total mass loss (no figure) but the front depths seem fit better. However, we need to notice that front depth from HYDRUS tends to slightly decrease in all range comparing to [Figure 2.6].

Actual data comparison also shows the improvement by the two-direction flow solution. The experimentally-based data from literature are used for this test as previously accomplished in

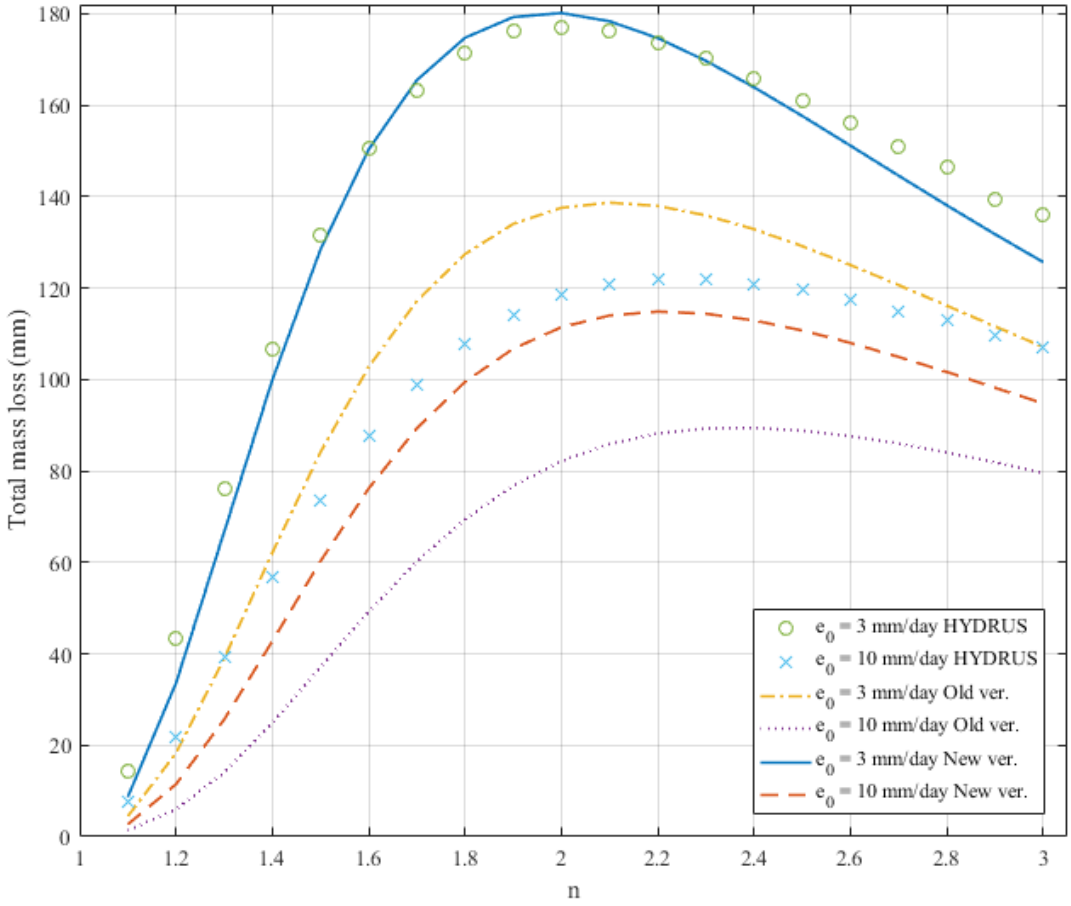

Figure 2.8: Plots of total mass loss with different soil properties for comparing results of HYDRUS and Eq. (2.10): Old version is with ($c = 1$) and new version is with ($c = 2$).

Sadeghi et al. (2012); Shokri and Salvucci (2011) [Figure 2.12]. These graphs, the relative evaporation rate versus the front depth normalized by the maximum depth Z_{max} , clearly show that evaporation rate is strongly related the front depth, and the activity of evaporation rate almost disappears when an evaporation depth is deeper than Z_{max} except for Chino clay. In the case of Chino clay, the significantly different behavior from the others may be caused by adapting lower potential evaporation due to the insufficient information (Sadeghi et al., 2012). The figure also shows differences between the old formulation ($c = 1$) and the proposed method ($c = 2$). The average relative evaporation rate is decreased by 89% in the old ones and 98% in the new ones after the

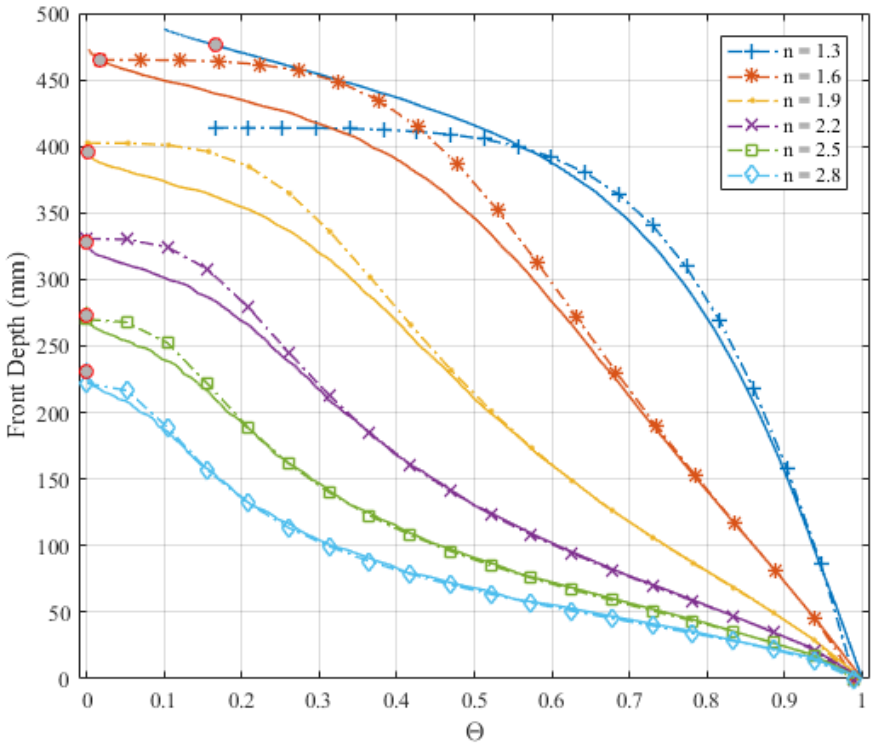

Figure 2.9: Dotted lines are new method and the lines are generated by HYDRUS (e_0 = 3mm/day). The filled circles indicates the end of Stage I (h=-1500kP).

stage change: these average values for Stage I were estimated about the relative depth 0.5 $\sim$ 1 (Z/Z_{max}), and between 1 and 1.5 for Stage II. These average values for Stage I are 0.722 for old methods and 0.657 for the new ones, and 0.077 and 0.011 for Stage II. Both graphs have similar behaviors as mentioned above, but a model taking into account two direction flows tends to result in more consistency at around the transition point where the Stage I is ending (the front depth reach the maximum value Z_{max}).

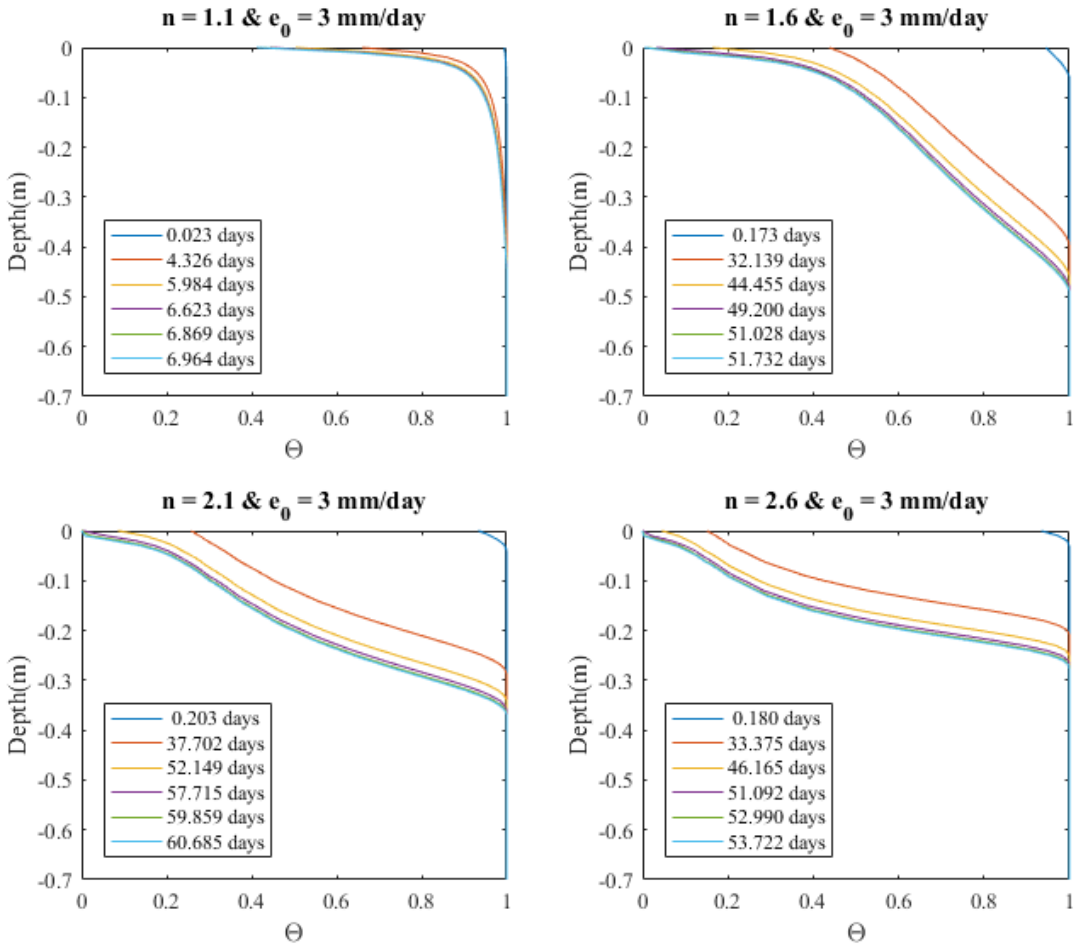

Figure 2.10: Detailed plots front depth behavior with different parameter 'n' in HYDRUS simulation.

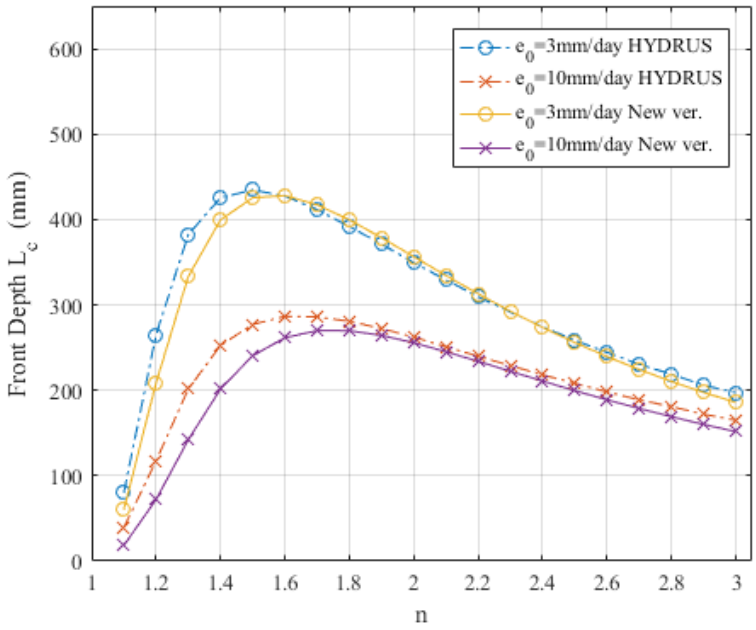

Figure 2.11: The same graph as [Figure 2.6] but the front depth is defined at the water content equal to 0.95 in results from HYDRUS. In [Figure 2.6], it is defined as 0.99.

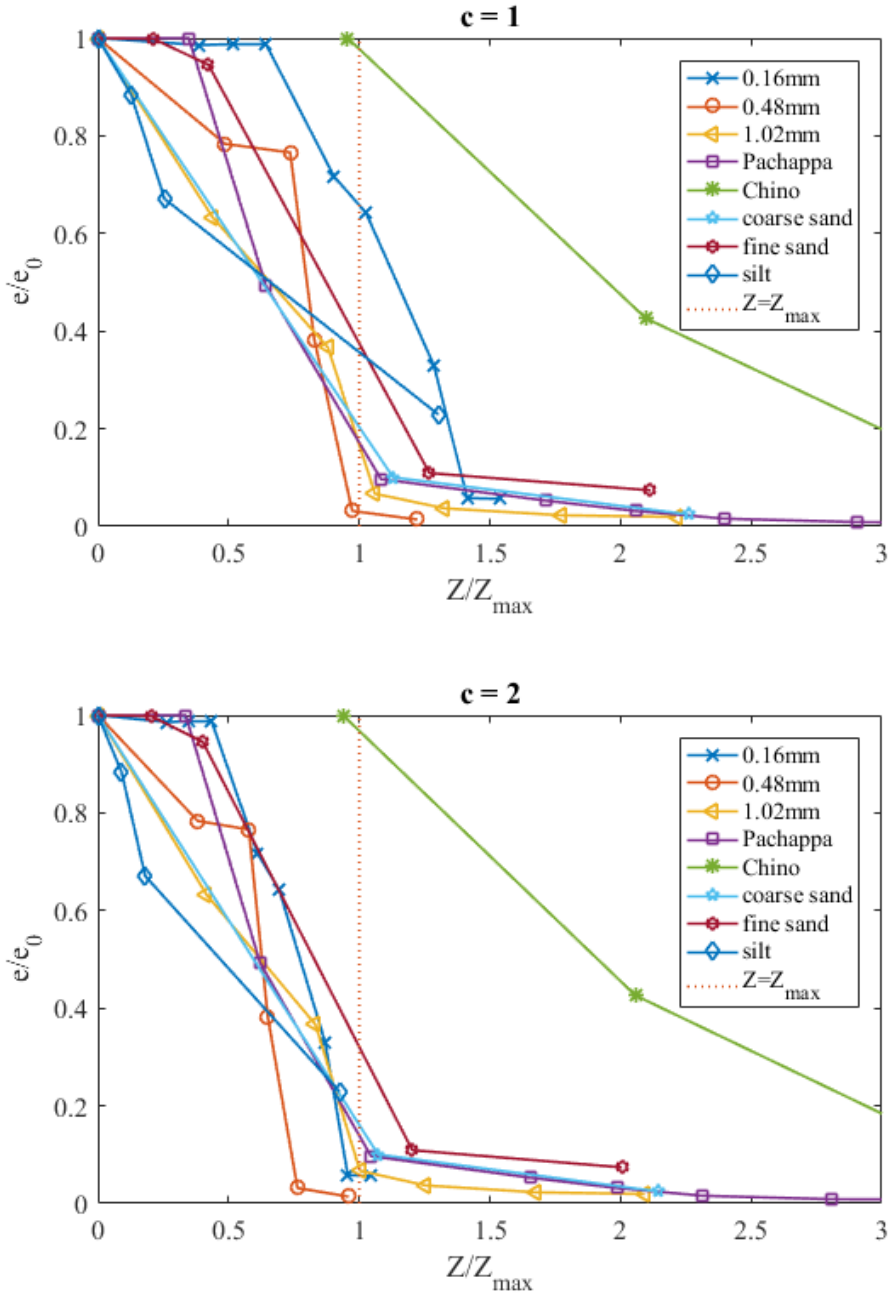

Figure 2.12: Experiment data plot using Z_{max} based on Eq. (2.18): it is the same as Sadeghi et al. (2012) if $c = 1$, and it is from the introduced method if $c = 2$.

2.5 Summary and Conclusion

This study introduces two-direction flow concept to catch both the improved prediction of the saturated depth and the conservation of mass. The change of key concept is explored by the inter-comparison between the previous model, the newly derived model, the mass-conservation model (HYDRUS1D), and observations. Different from the previous method (single-direction flow), the newly introduced model is derived through anisotropic Darcy's law but also Richards equation, which shows a need to consider both horizontal and vertical flows (hydraulic conductivities) to capture mass-conservation. The flows are finally caused by vertical pressure differences, but the water is supplied through two directions: vertical to vertical flow and horizontal to vertical flow within one grid cell. The new equation, which has an integral form, has a similar equation form as the previous method, so derivation of exact solution and simplified solution is possible as Sadeghi et al. (2012) has been done.

The previous model can mimic the trend between soil parameter 'n' and the front depth from HYDURS 1D, but its depth is significantly under estimated. The value of the total duration which is calculated from total masses Eq. (2.10) is not well predicted [Figure 2.5]. However, the newly introduced method Eq. (2.8) with ($c = 2$) results in better agreement [Figure 2.6; Figure 2.12] than with previous method [Figure 2.4; Figure 2.12], compared with "true" sets (e.g., the HYDRUS1D and experiment-based data from the literature). This result supports that the front depths can be controlled by a vertical direction flow, but they can be also significantly affected by horizontal direction flows. Adding horizontal flow within the front-depth equation, which then loses more water or has less resistance for the evaporation flux, gives such stronger agreement. In the consequences of cumulative mass Eq. (2.10) and total time Eq. (2.11), calculated based on the new method, they are well agreed with those by Richards equation [Figure 2.7; Figure 2.8]. Therefore, these results indicate that new theory gives an improvement of both the estimation of the maximum front depth and the mass conservation for a broad of soil category. Additionally, it shows a possibility of estimating the front depth not only for the end of Stage I but also for depending on different surface water contents [Figure 2.9].

Additionally, this study also highlights the advantage of analytical/semi-analytical solution for the saturated depth. Of course, A lot of computational time can be saved, but a confusion caused by numerical error can be also prevented. In the simulation of the HYDRUS 1D for Richards equation, the profiles of water contents with small value of 'n' makes difficult to identify the exact saturated depth, because the depth varies with different model settings. Due to this sensitivity, a criterion is required to avoid such errors. In this study, the locations of the front depth are extracted from the profiles at the water content equal to 0.99. This uncertainty may appear under measurement and it can consequently cause problems in the interpretation of experimental data and simulations. Therefore, further study is necessary to better understand the front depth and to develop appropriate criteria for it.

As previously mentioned, simulations with relatively small 'n' ($n < 2$) tend to have bigger discrepancy in [Figure 2.9]. It is still not clear whether there is an error in the numerical solution of the Richards equation or there are inherent difficulties with small 'n' in the new method. Also, one might question to the criteria of the pressure for the end of Stage I ($-1500kP$) for all soil types. Finally, the parameter ($c = 2$) in the new model could vary if there were anisotropy of hydraulic conductivity, as might be normally found in a real soil system. It is necessary to investigate these unresolved problems in future research, to understand better soil-water movement and to improve its prediction.

3. INVESTIGATION OF DIFFUSION-BASED SOIL EVAPORATION MODEL FOR HETEROGENEOUS PORE SIZES AND ITS APPLICATION WITH SOIL RESISTANCE SCHEME

3.1 Introduction

As mentioned in previous section, some evaporation studies utilize a scaled (relative) evaporation rate e/e_0 (Lehmann, 2008; Shahraeeni et al., 2012; Haghighi et al., 2013) also known as a beta function β (Kondo et al., 1990; Sakaguchi and Zeng, 2009), where e is actual evaporation rate and e_0 is potential evaporation rate. This scaled drying rate is useful to identify major controls beyond the atmosphere condition, and it tends to have several stages responding to the different main controls of soil evaporation process. Starting from a fully saturated condition, the drying rate remains constant as much as a potential drying rate (Stage I), until the surface water contents reaches 'critical water content' or often known as 'field capacity' (Sherwood, 1930; Sakaguchi and Zeng, 2009). Then, the evaporation rate decreases and this period is called Stage II. The decreasing rate becomes slow, and eventually it can be called as Stage III. Based on this concept, the features of relative evaporation are defined as two or three stages (Idso, 1974; Lehmann, 2008), and this behavior is often parametrized resting on empirical methods (Merlin et al., 2016; Chanzy and Bruckler, 1993; Kondo et al., 1990; Sakaguchi and Zeng, 2009).

Several attempts have been made to capture such feature of the relative evaporation based on physical-based relationships (e.g., diffusion theory) within Viscous Boundary Layer (VBL) (Suzuki and Maeda, 1968; Schlünder, 1988). S&M (Suzuki and Maeda, 1968) and SCH (Schlünder, 1988) models follow the critical behavior of soil evaporation, where the relative evaporation rate decreases slowly at the beginning (not a perfectly constant) and then starts to sharply drop at some points during a decrease in the water content. Such concept makes models possible to represent Stage I and Stage II in one physical equation. Here, this behavior varies depending on the wetted pore size P as in SCH or the particle size W (total or maximum pore size including the

wetted pore) in S&M, if viscous layer thickness δ is fixed. The division of Stage I and Stage II more clearly appears in the case of small values of W or P [Figure 3.1]. In such models, evaporative flux from the soil surface is mainly regulated by (1) thickness of wind layer and (2) the relative wet-pore size (wet area). Based on these features, many studies have been accomplished such as Lehmann et al. (2018); Shokri et al. (2008); Shahraeeni et al. (2012); Haghighi et al. (2013). Inside the viscous boundary layer, advection and diffusion co-exist, and among them the diffusion process is dominant for the bare soil evaporation (Haghighi et al., 2013). However, the model in VBL does not always successfully capture all possible evaporation activities. Sometimes experimental data show opposite or different behaviors: Stage I can be very short with small P and W (Chanzy and Bruckler, 1993; Teng et al., 2014). Therefore, it is still necessary to update such models via more precise understanding of such phenomenon.

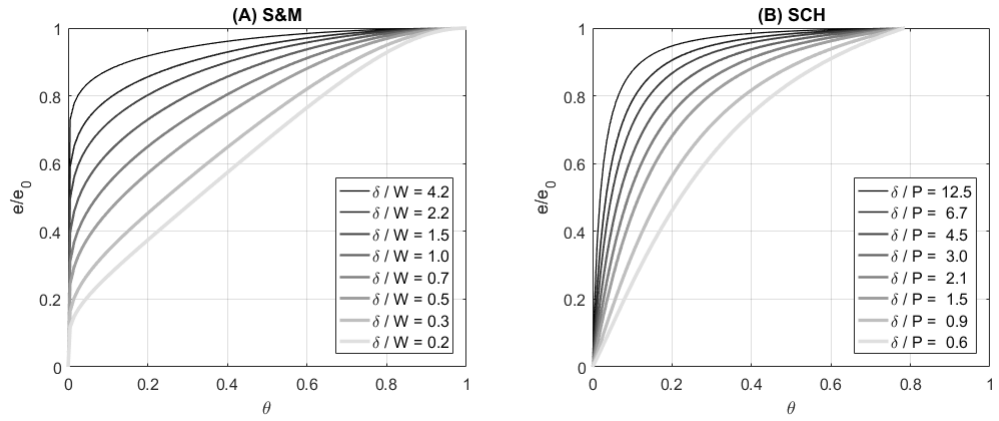

Figure 3.1: Plots of S&M models and SCH model (Suzuki and Maeda, 1968; Schlünder, 1988).

Considering only VBL cannot be sufficient to demonstrate evaporation activities, because the evaporating water is supplied through the Soil Layer (SL) in a liquid or vapor state. Haghighi et al. (2013) shows the evaporation rate is controlled by the condition within VBL including atmospheric demand and surface water contents, but its water supply is also controlled by soil resistance. The study indicates that conditions of both below and above soils need to be taken into account in soil

evaporation models. Applying the resistance terms based on both the Schlünder model for VBL and Darcy's law for SL, the combined models can be successfully simulated compared to experimental data, and it alleviates the weakness of the only diffusion-based model (e.g., overestimation in stage II and III). Shahraeeni et al. (2012) performed a simulation utilizing both S&M model for VBL and the front depth theory for SL (Lehmann, 2008) to reflect the additional contribution to the evaporation through invaded pores (from deep soil). The simulated results follow well the experimental data in several ambient conditions with relatively homogeneous soil (small variance of pore size distribution). These studies highlight that the role of soil is important as much as the atmospheric condition above the soil surface.

In spite of the consistent model improvement, an evaporation model cannot satisfactorily mimic observations yet, and some significant disagreement still exists against gauged data in some environment conditions. The reason of such errors is not clear, and so, investigation is still necessary to reduce the gap between our understanding and the complicated natural phenomenon. Some experiments show the Stage I rapidly disappears under high demand of atmospheric condition (Shahraeeni et al., 2012; Mosthaf, 2014). Although the diffusion based models (SCH or S&M) with soil resistance models are used as a main controll on Stage I or II, these simulations are not able to adequately fit to experimental data (Mosthaf, 2014). Mosthaf (2014) tested numerical simulation based on Schlünder's model considering key elements (e.g., thermal conductivity, vapor/liquid flux, etc.) inside of the soil system, and the sensitivities and behaviors were explored applying different simulation parameters. In the one of the investigations, Mosthaf (2014) cast doubt on the selection of pore sizes for the SCH model (usually mean or median of pore size are used for the model), because pore sizes of natural soil are not actually homogeneous. From this idea, a calculation of the pore size is attempted by Mosthaf (2014) in Representative Elementary Volume (REV) scale, based on Young-Laplace equation to apply its variation into the Schlünder's models. The test results in some improvement of the model at high demand atmosphere condition but also shows a needs of more study.

However, the effect of pore-size heterogeneity, which results in ever-changing wet-pore size

with surface water contents, is not sufficiently investigated yet at REV-scale with various soil types. Like as Mosthaf (2014), the equations are based on assumptions such as homogeneous pore (SCH) or continuum system (S&M), so it may not be adequate to simulate relative evaporation due to their over-simplification (Mosthaf, 2014). Moreover, most of the experimental data and simulations have had relatively large particles and a large Van Genuchen (van Genuchten, 1980) parameter n which means the soil is close having homogeneous conditions. Therefore, it is necessary to investigate VBL-related models in the case of relatively inhomogeneous soil grains.

Moreover, relative evaporation models in VBL often employ S&M and SCH model to connect between soil surface and the atmosphere. Although the two are derived based on diffusion-based theory and both seem to have similar behaviors, their outputs are not close to each other (Haghighi et al., 2013). Therefore, it is necessary to revisit the two models and check their dissimilarity.

In this study, (Section 3.2) diffusion-based models, S&M and SCH, are explored to demonstrate their similarities and to show the proper use of the models for soil evaporation. Then, (Section 3.3) inhomogeneous model (modified SCH) is introduced based on Young-Laplace equation and pore size distribution. (Section 3.4) The new model is tested through a soil-atmosphere resistance scheme (Haghighi et al., 2013), to identify their characteristics about a wide range of soil properties. Applying such heterogeneously wetted soil pores, this investigation will give a great insight of its role in between atmosphere and soil system.

3.2 Two Homogeneous Diffusion Models

In this section, the two diffusion-base models are briefly introduced and then investigated to prove their similarity. S&M is founded on the well-known diffusion-equation (Fick's law), but it needs to be calculated using a numerical methods. On the other hand, SCH model are an asymptotic solution with $\frac{\text{length of mean free path}}{\text{radius of wet patches}} \rightarrow 0$, but its simulation is very straightforward. As mentioned before, the both are based on diffusion theory and they seem to have similar behaviors, but their results are not close to each other [Figure 3.1].

3.2.1 S&M Model

Suzuki and Maeda (1968) solve the diffusion equation of Fick's law in 2D space, with a horizontal direction (x) and vertical direction (z) [Figure 3.2]. The two-dimensional diffusion-equation

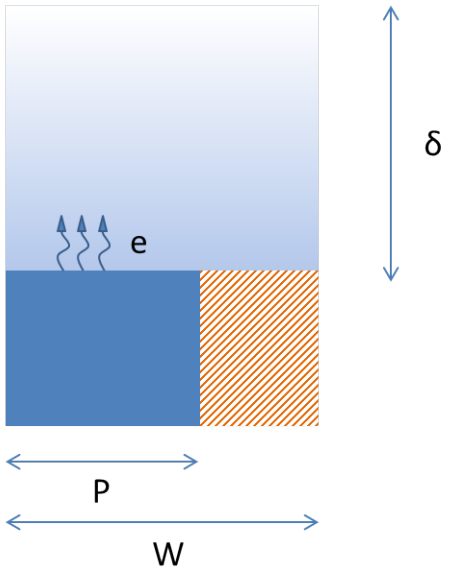

Figure 3.2: The concept of S&M models: P is a wetted part for evaporation e, W is a width of wet and dry surface, and δ is viscous boundry layer which is usually determined by windspeed. SCH model, on the other hand, consider 2-D circle for P and 2-D square for W.

is

$$\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial z^2} = 0 \tag{3.1}$$

where C is concentration. With boundary conditions including surface water contents $\theta_{surf} = P/W$ (see Suzuki and Maeda (1968)), the solution for the relative evaporation model depends on 3 variables as Eq. (3.2).

$$\frac{k}{k_0} = S\&M(\theta_{surf}, \delta, W) \tag{3.2}$$

where k is actual flux (conductivity m/s), k_0 is potential flux, θ_{surf} is 1D surface water content, δ is the thickness of the viscous boundary layer, W is the width of a unit surface including soil

particles and pores which is often regarded as being proportional to particle size (Shokri et al., 2008), P is wet pore size. The solutions are introduced by Suzuki and Maeda (1968) as

$$\frac{k}{k_0} = \frac{I(\aleph)I'(\beth)}{I'(\aleph)I(\beth)}, \qquad \frac{\delta}{W} = \frac{I(\aleph)}{2I'(\aleph)}, \qquad \theta_{surf} = \frac{sn^{-1}\left(\beth'/\aleph', \aleph'\right)}{I'(\aleph)} \tag{3.3}$$

where $\aleph$ and $\beth$ are moduli(parameters) which have a geometric space and are only used internally, I is the complete elliptic integral of the first kind, $I'(\beth)$ is the same as $I(\sqrt{1-\beth^2})$, $\beth'$ is the same as $\sqrt{1-\beth^2}$, and sn is the Jacobian elliptic function. We need to note that this solution requires additional numerical methods (e.g., optimization) to get the actual numbers.

3.2.2 SCH Model

For SCH model (Schlünder, 1988), the solution is derived based on Stefan diffusion and Knudsen diffusion theories. Using ratio of the free path and a pore size, an asymptotic solution is introduced as

$$\frac{k}{k_0} = SCH(\theta_{surf}, \delta, P) = 1/\left[1 + \frac{P}{\delta}\frac{2}{\pi}\sqrt{\frac{1}{4\theta_{surf}}}\left(\sqrt{\frac{\pi}{4\theta_{surf}}} - 1\right)\right] \tag{3.4}$$

where P is pore size and θ_{surf} is equal to $(\frac{P}{W})^2\frac{\pi}{4}$ which represents 2D surface water content. To express those θ_{surf} in 3 dimensional way, they become $\theta_{3D} = (\frac{P}{W})^2\frac{\pi}{4}\frac{L}{L}$ for SCH and $\theta_{3D} = \frac{P}{W}\frac{W}{W}\frac{L}{L}$ for S&M, where L is a vertical length (Haghighi et al., 2013).

3.2.3 The Proof of Similarity

To understand each model and to identify the reason for their different results [Figure 3.1], we need to reconsider the meaning of each parameter carefully. Parameter W is normally regraded as a total width or particle size, P is a pore size, and their relationship can be approximately expressed as $W \approx 3P$ (Haghighi et al., 2013; Gupta and Larson, 1979). If the soil surface was homogeneous, P represents a mean pore size and W is a mean particle size. In a homogeneous soil system, even though the water content on the surface is decreased by evaporation, pore sizes which are active or inactive (whether currently contributing evaporation rates) cannot change. In

other words, while surface water contents change, averaged pore size should be fixed. The water content is altered by the transition of each pore state (active↔inactive). If active pore sizes vary according to the change of water content during the simulation with S&M or SCH, they cannot correctly demonstrate a evaporation rate.

Therefore, the pore sizes for those models need to be a constant in homogeneous soil system, and it is proper to carry out simulations with assumption that the width of dry surface W changes along with θ_{surf}.This idea agrees with Shahraeeni et al. (2012) and we can conclude that the Schlünder's approach is more plausible to construct a relationship between surface water content and relative evaporation rate about soil surfaces. If this idea is applied to S&M with assumptions that SCH surface water content $\theta_{surf} = \theta_{3D} = (\frac{P}{W})^2 \frac{\pi}{4} \frac{L}{L}$ and S&M surface water content $\theta_{3D} = \frac{P}{W} \frac{W}{W} \frac{L}{L}$ are the same, W varies as $P/\sqrt{\frac{4}{\pi}\theta_{surf}}$. Then, the S&M equation can be rewritten as

$$\frac{k}{k_0} = S\&M\left(\theta_{surf}, \delta, P/\sqrt{\frac{4}{\pi}\theta_{surf}}\right) \tag{3.5}$$

The results of test simulations from modified-S&M Eq. (3.5) and Schlünder's equations Eq. (3.4) are not the perfectly same but reasonably well matched [Figure 3.3]. Through this simple test, we can conclude that the two methods are close to each other when the concepts of surface water contents are identical, and the result reveals that SCH methods are a reasonable way to achieve both the simulation of evaporation rate and reduction of computational time.

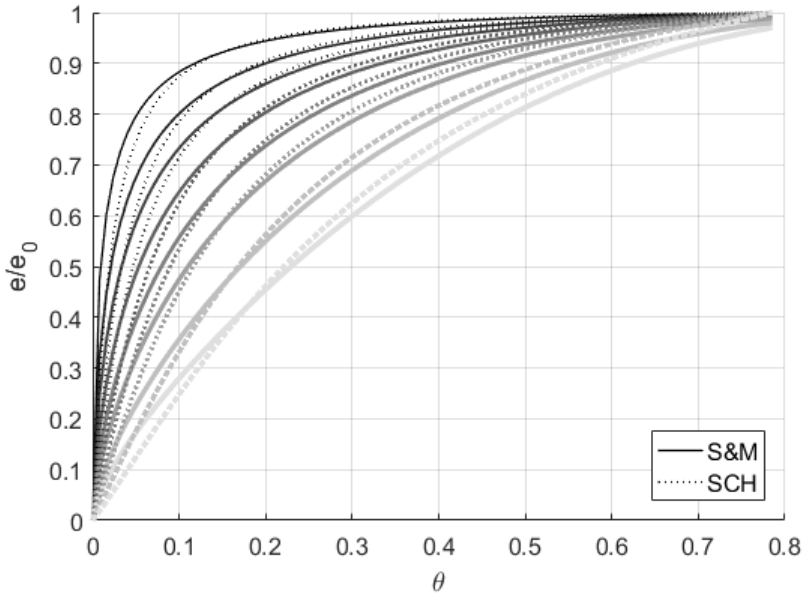

Figure 3.3: Comparison between SCH Eq. (3.4) and modified S&M Eq. (3.5)

3.3 Diffusion Model for Heterogeneous Soil Grain

The basic idea used to apply the model to heterogeneous soil pores is that the representative pore changes following the state of surface water content. Through this idea and derivation, this application will lead the pore size and the water contents to have a similar behavior with respect to unsaturated hydraulic conductivity. As shown in the exploration in the previous section, three input variables for S&M or SCH are required to simulate relative evaporation within the surface boundary layer. Among the parameters, two of them are directly related to soil properties, which indicates that the soil heterogeneity can affect the simulation performance. At a surface with unsaturated conditions, due to that the low effect of gravitational force, the force for water suction at the surface becomes dominant, mostly caused by matric potential or primarily capillary forces (Lehmann, 2008). When the surface water content decreases within a heterogeneous soil system, the relatively small pores tend to be the main contributor of evaporation [Figure 3.4]. On the other hand, large pores tend to dry due to relatively small negative pressures: the force of suction is relatively lower and gravitational force is dominant).

Therefore, The evaporation model needs to apply a more realistic activity of the interface between VBL and SL. [Figure 3.4]. Ideally, the mean pore sizes contributing to the evaporation (see S&M or SCH model) could vary along with the alteration of the surface water content (Mosthaf, 2014). Second, volumetric water content for usual soil models is based on 3 dimensional spaces ($m^3 \cdot m^{-3}$). The effective water contents, which just face to the atmospheric boundary layer in 2-dimensional way, may not be the same as the 3-dimensional water contents. Here, we would like to investigate how the application of non-homogeneous soil surface affects the result of relative evaporation models. In this section, effective pore sizes and effective water contents are introduced and the behaviors are explored.

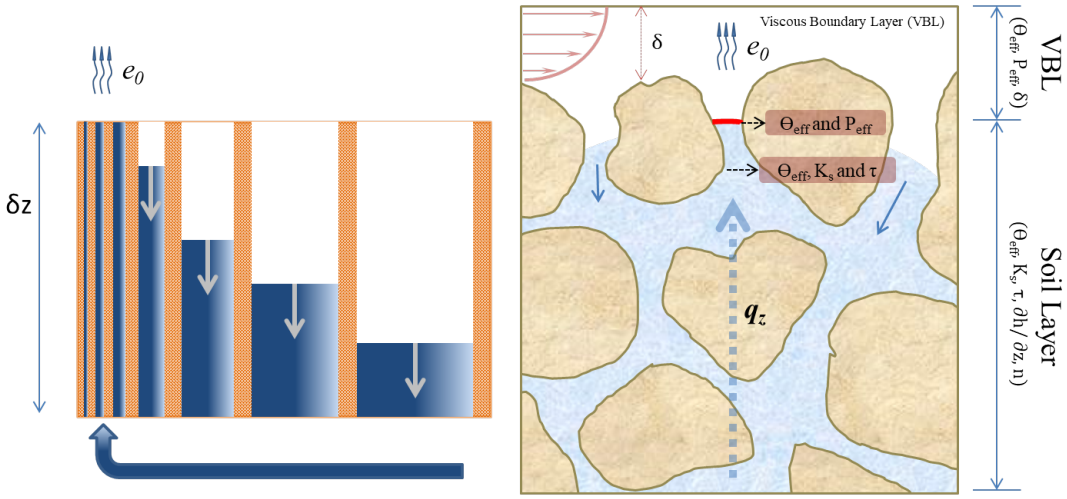

Figure 3.4: Scheme of the water behavior at a surface. Within Viscous Boundary Layer (VBL), k/k_0 is the function of $\Theta_{eff(2D)}$, P_{eff}, and δ which are the effective surface water contents, the effective pore size, and the thickness of VBL. In the soil layer, the vertical flow q_z follows Darcy's law, which consists of $\Theta_{eff(2D)}$, τ, K_s, $\partial h/\partial z$, and n, where τ is the tortuosity and n is a parameter for soil feature if the Van Genuchten (VG) retention curve (van Genuchten, 1980) is used.

3.3.1 Assumptions and Basic Equations

For this study, the water movement of soil system is conceptualized as capillary flows (by capillary potential). The matric potential is a better term than capillary potential which is a part of

the matric potential in soil system (Hillel, 2003). However, such a concept is reasonable because capillary force is a main driver and widely applied due to its simplicity in various soil related studies such as unsaturated hydraulic conductivity model (Mualem, 1976).

Here we assume that the water holding in a soil system is mainly governed by capillary force, which makes it possible to construct a relationship between the pore sizes and pressure heads using Young-Laplace equation as Mosthaf (2014) and Lehmann (2008). The Young-Laplace equation is

$$h = \frac{2\sigma}{\rho g R} = \frac{4\sigma}{\rho g P} \tag{3.6}$$

where g is the gravitational acceleration, ρ is the water density, h is the capillary pressure head, σ is the surface tension ($8.5 \times 10^{-3}\ kg \cdot m^2 \cdot s^{-2}$) (Tschapek et al., 1978), P is the diameter of the pore, and R is the radius of the pore. A contact angle (it is not presented in Eq. (3.6)) should be included but for relatively small pore size ($R < 300\mu m$) it is nearly constant and usually set to one (Lehmann, 2008; Li et al., 2014). We also assume that surface tension σ is the same among any pore, and can be also canceled out in the further steps of derivations. With the corresponding pressure h_m, the mode or average value of pore size distribution P_m is assigned as a base line, which can be estimated via a retention curve equation. Using Eq. (3.6), any pore size P can be calculated using base line values (h_m,P_m) as

$$P = \frac{h_m}{h} P_m \tag{3.7}$$

We need to note that $h_m \cdot P_m$ is a constant based on Eq. (3.6), so only one of the two needs to be known. However, they represent actual soil properties, and these values can be better to get independently through experiments. For instance, h_m can be regarded as a the air entry pressure h_g or estimated from pore-size distribution (retention curve). P_m can be earned from direct soil test, or it can be predicted using h_m with Eq. (3.6).

In this study, to connect between pressure heads and water contents, the Van Genuchten (VG)

retention curve (van Genuchten, 1980) is applied for this study.

$$\Theta = \frac{\theta(h) - \theta_r}{\phi - \theta_r} = [1 + (\alpha h)^n]^{-m} \tag{3.8}$$

where Θ is a normalized water content, θ is a volumetric water content, θ_r is a residual water content, ϕ is a saturated water content or porosity, n is a parameter related to a pore size distribution, m is $1 - 1/n$ for this study, and α is equal to $1/h_g$. If h_m corresponds to the mode of pore size, the pressure head from VG model is

$$h_m = h_g \left(\frac{n-1}{mn+1} \right)^{1/n} \tag{3.9}$$

If h_m was set to an average value, air entry pressure h_g becomes close to h_m, and this choice gives a further derivation simplification. However, existing retention curves such as VG models tend to provide unrealistic pore sizes (because of the long tail of the distribution) at the very wet/dry condition, and consequently this leads mean pore sizes to be very small in some cases, so the mode of pore size is used for P_m with h_m for this study. If pore size P_m was given, it is primarily used as mentioned previously. Using these simple capillary-based relationships, effective pore size and surface water content are derived in the next section.

3.3.2 Effective Pore Size

As in [Figure 3.4], the effective/active pore sizes consist of relatively smaller pores than the mean of total pore size $\bar{P}$, because relatively thin pores may be main providers of surface evaporation under an unsaturated state. The water held in relatively large pores might flows out in gravitational and horizontal way, furnishing the active pores (small pores) as the evaporation (Shahraeeni et al., 2012). From this idea, we derive the mean of this active pore sizes; it will be called here the effective pore size. With pore size distribution $d\Theta/dh$ and Eq. (3.7), the effective pore size P_{eff}

can be calculated as

$$P_{eff} = \bar{P} = \frac{\int_{h_{surf}}^{-\inf} \frac{d\Theta}{dh} \cdot P dh}{\int_{h_{surf}}^{-\inf} \frac{d\Theta}{dh} dh} = h_m P_m \frac{\int_{\Theta_{surf}}^{0} \frac{1}{h} d\Theta}{\int_{\Theta_{surf}}^{0} d\Theta} = P_m \left(\frac{h_m}{h_g}\right) \frac{\left[1 - \left(1 - \Theta_{surf}^{1/m}\right)^m\right]}{\Theta_{surf}} \quad (3.10)$$

where the mean pore size $\bar{P}$ becomes the same as the effective pore size P_{eff} when $\bar{P}$ is no longer a constant, $\Theta_{surf}\ [m^3 \cdot m^{-3}]$ is normalized surface water contents, and h_{surf} is the surface pressure head. One may notice that the derivation is very similar to the Mualem models (Mualem, 1976) and the Eq. (3.10) is tested with different parameters [Figure 3.5]. Simulation of this equation indicates that the effective pore size is always smaller than P_m (h_m/h_g) as expected. In other words, when the surface water content decreases (drying), the decreasing rate (Stage I $\rightarrow$ Stage II) of relative evaporation becomes more moderated than those in homogeneous model (smaller P in Eq. (3.4)), due to smaller pore has a longer Stage I. Another characteristic is the SCH model with the effective pore sizes becomes additionally depending on soil parameters n and h_g. If n is large, which means the pore size distribution is more homogeneous, the simulation is close to the original SCH model.

3.3.3 Effective Surface Water Content

Like as the effective pore sizes, relatively thin pores also affect surface water contents [Figure 3.4]. The effective surface water contents, which refer to 2-dimensional area or contact area to the atmosphere, may not be the same as the volumetric water content (3D). Therefore, the area ratios between water filled pores and total pores can be likewise calculated based on the Young-Laplace equation and the pore size distribution. The relationship is defined as

$$\Theta_{eff(2D)} = \left[\frac{P_{wet}}{W}\right]^2 = \left[\frac{\int_{\Theta_{surf}}^{0} \frac{1}{h} d\Theta}{\int_{1}^{0} \frac{1}{h} d\Theta}\right]^2 = \left[1 - \left(1 - \Theta_{surf}^{1/m}\right)^m\right]^2 \quad (3.11)$$

where $\Theta_{eff(2D)}$ is a normalized water content tangent to the atmosphere or an area ratio between active pores P_{wet} and total pores W. It is possible to apply an area-based ratio A_P/A_W instead of

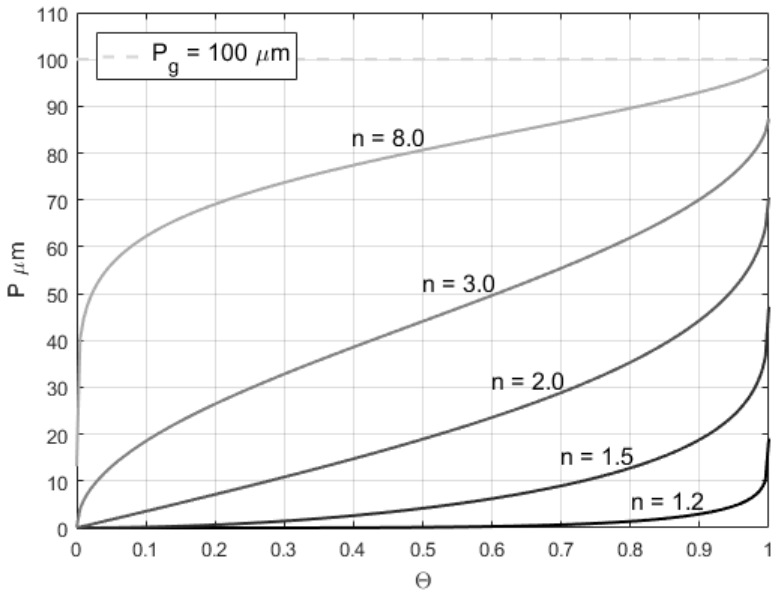

Figure 3.5: The behavior of effective pore size based on Eq. (3.10). Mean of pore size is set to 100 μ m, the parameters n are 1.2, 1.5, 2, 3, 8 and $m = 1 - 1/n$.

P_{wet}/W for Eq. (3.11) but it does not provide a closed-form.

The derivation of this equation is identical to Van Genuchten-Mualem (VGM) (van Genuchten, 1980; Mualem, 1976) except a tortuosity parameter and a saturated hydraulic conductivity: VGM can be written as $K(\Theta) = K_s\tau^{0.5}\Theta_{eff(2D)}$. This VGM model is widely applied and validated in soil-related simulations, which indicates these derived parameters (e.g., effective pore size and water contents) have a similar role at the interface. The driver for the evaporation in VBL becomes $\Theta_{eff(2D)}$, P_{eff}, n, and δ at the soil surface (the red line in Figure 3.4). Similar to the parameters of the interface for VBL, the key variables for the vertical flow q_z at just below the soil surface are τ, K_s, $\partial h/\partial z$, and n including $\Theta_{eff(2D)}$ too in VGM [Figure 3.4]. The expression of VGM represents the active proportion of 2-D area as the cross section of water pipe passing through between soil layers and the surface layer. In the case of Eq. (3.11), it can be seen that the effective water contents $\Theta_{eff(2D)}$ represent the wetted-pore area facing to the atmosphere and contributing actual evaporation toward the atmosphere. To get its surface water content $\theta_{eff(2D)}$, we assume that residual water content is close to zero so it is negligible. Then, from the 3-D porosity ϕ, the

two-dimensional water content can be calculated as

$$\theta_{eff} = \frac{\pi}{4}\left[\frac{P_{wet}}{W}\phi^{1/3}\right]^2 = \frac{\pi}{4}\Theta_{eff(2D)}\phi^{2/3} \tag{3.12}$$

where $\pi/4$ is for the circle shape of pore, and $\phi^{1/3}$ is one dimensional porosity to get the total wet-able surface area $[W/\phi^{1/3}]^2$. The plots for the example are displayed in [Figure 3.6], which applied both effective pore size and effective water contents in the SCH model. Although effective pore size tends to make Stage I longer, relative evaporation rate eventually decreases by effective water contents.

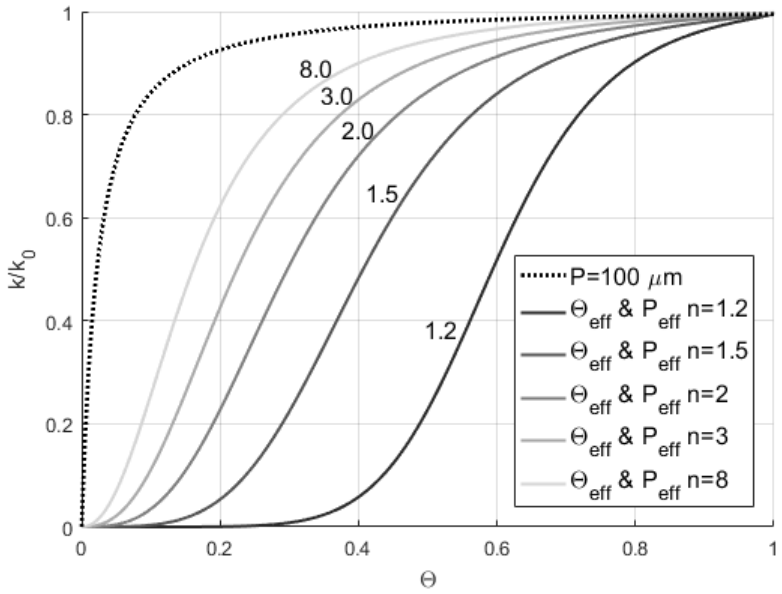

Figure 3.6: Simulation using SCH model using the effective pore size and surface water content (lines). Average pore size 100 μm and δ=0.0023m are used here. Original SCH model is depicted for the reference (dotted line).

3.4 Setup for Model Test

In this section, the model of bare soil evaporation is derived applying non-homogeneous soil grains through the effective pore size and effective water contents, based on Schlünder (SCH)

models and a soil resistance scheme like as Haghighi et al. (2013). The following part explains a simulation setup for a test including required parameters and observations to compare.

3.4.1 Modified-SCH Model

If both the effective pore size Eq. (3.10) and effective surface water content Eq. (3.11, 3.12) are applied in the SCH model, Eq. (3.4) becomes the modified-SCH model:

$$\frac{k}{k_0} = \frac{e}{e_0} = SCH[\theta_{eff}\left(\theta_{surf}, n\right), \delta, P_{eff}\left(\theta_{surf}, P_m, n\right)] \tag{3.13}$$

where k/k_0 is the ratio of mass transfer, e/e_0 is in evaporation term, $\theta_{eff}\,[m^2 \cdot m^{-2}]$ is 2D surface water contents, $\theta_{surf}\,[m^3 \cdot m^{-3}]$ is the volumetric water content at the surface, and P_m is the mode of pore size [Figure 3.6]. The modified-SCH (MSCH) model depends on four parameters that are the volumetric water content at the surface θ_{surf}, the thickness of viscous boundary layer δ, the mode of pore size P_m, and the parameter of pore size distribution n. In most soil studies, some parameters are given, but P_m and δ (wind speed) are not generally provided. A value for pore size can be analogized through soil types or approximately estimated though a simple calculation via grain sizes ($W_{averege}/3$) if they are provided. It is also possible to estimate pore size through the Young-Laplace Equation Eq. (3.6, 3.7).

3.4.2 Full Resistance Scheme

Haghighi et al. (2013) utilize total resistance r_t, which consists of a resistance inside of the soil r_s and within boundary-layer between the surface and the atmosphere (VBL) r_{BL}.

$$r_t = r_{BL} + r_s\,[s \cdot m^{-1}] \tag{3.14}$$

$$r_{BL} = \frac{1}{\left(\frac{k}{k_0}\right)_{SCH}} \frac{1}{k_0} = \frac{1}{\left(\frac{k}{k_0}\right)_{SCH} \cdot \frac{D}{\delta_m}} \tag{3.15}$$

$$r_s = \frac{C_s - C_\infty}{\rho_w K_{eff} \frac{\partial h}{\partial z}} = \frac{p_s - p_\infty}{\rho_w 4K(\theta_{surf})} \frac{M_m}{RT} \tag{3.16}$$

where $C_s - C_\infty$ is the difference of vapor concentration, $p_s - p_\infty$ is the difference of the vapor pressure, M_m (≈0.018 $kg \cdot mol^{-1}$) is the molar mass of water, R (≈8.314 $J \cdot mol^{-1}$) is the universal gas constant, ρ_w (≈1000 $kg \cdot m^{-3}$) is density of water, $\frac{\partial h}{\partial z}$ is set to one like as Haghighi et al. (2013), T is absolute temperature, D is the diffusion coefficient, δ_m is the height of viscous boundary layer, and K_{eff} is $4K(\theta_{surf})$ (Haghighi et al., 2013).

Based on the total resistance model introduced above, due to the fact that $k = 1/r_t$ and $k_0 = D/\delta_m$, the relative evaporation can be expressed as

$$\frac{k}{k_0} = \frac{1}{r_t D/\delta_m} = 1/\left[\frac{1}{\left(\frac{k}{k_0}\right)_{SCH}} + \frac{p_s - p_\infty}{\rho_w 4K(\theta_{surf})}\frac{M_m}{RT}\frac{D}{\delta_m}\right] \tag{3.17}$$

If rearrange it by applying the heterogeneity of soil pore (the effective pore size and the effective water contents), it becomes

$$\frac{k}{k_0} = \frac{1}{r_t D/\delta_m} = 1/\left[1 + \frac{P_{eff}}{\delta_m}\frac{2}{\pi}\sqrt{\frac{1}{4\theta_{eff}}}\left(\sqrt{\frac{\pi}{4\theta_{eff}}} - 1\right) + \frac{p_s - p_\infty}{\rho_w 4K(\theta_{surf})}\frac{M_m}{RT}\frac{D}{\delta_m}\right] \tag{3.18}$$

We need to note that this final equation reflects three important layers, which are VBL, SL, and their interface. The model takes soil pore heterogeneity of SL and the interface into account, which finally affects soil-atmosphere interaction (evaporation) in the VBL.

A relationship for viscous boundary layer δ_m is used: (Lehmann, 2008; Haghighi et al., 2013; Haghighi and Or, 2013)

$$\delta_m = 2.26 \cdot U_\infty^{-1/2}\ [mm] \tag{3.19}$$

where U_∞ is an air velocity. This equation is also validated under turbulent conditions (Mosthaf, 2014). The vapor pressure and vapor concentration difference are also obtained based on Clausius-Clapeyron relation as in Haghighi and Or (2013)

$$p(T) = 611 \cdot \exp\left(\frac{\lambda M_m}{R}\left(\frac{1}{273} - \frac{1}{T}\right)\right)\ [Pa] \tag{3.20}$$

$$p_s - p_\infty = p(T_s) - RH \cdot p(T_\infty)\ [Pa] \tag{3.21}$$

where $\lambda \approx (2450 \cdot 10^3\ J \cdot kg^{-1})$ is the latent heat of vaporization, $T_s[K]$ is the surface temperature, and $T_\infty[K]$ is the atmosphere temperature. Last, the diffusion coefficient D and saturated hydraulic conductivity K_s based on the Kozeny-Carman method are also calculated as in Haghighi and Or (2013) and Haghighi et al. (2013)

$$D = 1.49 \cdot 10^{-7} \cdot T_\infty - 1.96 \cdot 10^{-5}\ [m^2 \cdot s^{-1}] \tag{3.22}$$

$$K_s = \frac{(3P)^2 \phi^3 g}{180(1-\phi)^2 \nu}\ [m \cdot s^{-1}] \tag{3.23}$$

where P is a mean pore size, ϕ is porosity, g (≈9.81 $m \cdot s^{-2}$) is the acceleration of the gravity and ν ($\approx 10^{-6}\ m^2 \cdot s^{-1}$) is the kinematic viscosity. For unsaturated hydraulic conductivities, Van Genuchten-Mualem models (van Genuchten, 1980; Mualem, 1976) are applied.

3.4.3 Simulation Setup

To explore the simulation of Eq. (3.17) or Eq. (3.18), the results are compared with several experimental data from literatures (Kondo et al., 1990; Chanzy and Bruckler, 1993; Griend and Owe, 1994; Aluwihare and Watanabe, 2003; Lehmann et al., 2018; Merlin et al., 2016). The majority of the data do not provide sufficient information including meteorological variables and soil properties, so some of parameters (e.g., temperature) are approximately selected or estimated through the other subsidiary information such as graphs. However, the soil parameters which are important in this study are selected as described below. The mode of pore size are estimated based on the VGM parameter 'n'. Additionally, as in a previous study by Lehmann (2008), air entry pressure (h_g or α) is predicted via a model Eq. (3.24) fitted by Carsel and Parrish (1988) as

$$\alpha = 0.0087(n-1)\ [mm^{-1}] \tag{3.24}$$

which is also used for the calculation of the mode of pore size distribution Eq. (3.6) and Eq. (3.9). If VGM parameter 'n' was not given, it is approximated through soil categories.

Aluwihare and Watanabe (2003) conducted an experiment using Toyoura Sand with 0.445 porosity, 0.018 cm particle size (this value is not used here), and 4.55 for n. Due to the fact that the actual velocity is not provided, wind speed was assumed to be 0.4 m/s for the simulation: the experimental instrument can control wind velocities up to a maximum value of 0.7 m/s (Aluwihare and Watanabe, 2003) [Figure 3.7]. The relative humidity (RH) is set as 0.7 through their graph, and surface and air temperature are 30 C°. In the case of Griend and Owe (1994) [Figure 3.7], surface resistance is studied for fine Sandy-Loam under field condition which has T_s=27.8 C°, T_a=32.8 C° and RH = 0.5. The wind speed is up to 4 m/s but wind speed is also unknown (3 m/s is used here). Other parameters are set as n=2.5, and θ_s = 0.4 for the simulation in [Figure 3.7]. Other surface resistance vlaues for loam and sand are gauged and calculated by Kondo et al. (1990), and the experiment conditions are air velocity = 1 m/s, T_s=35 C°, and T_a=35 C°. The porosities of the loam and sand are 0.39 and 0.49. The un-provided parameters are set as n = 1.7 for loam, 2.5 for sand, and RH = 0.5 for the both [Figure 3.8]. Last, the relative evaporation rate is calculated through an empirical method by Chanzy and Bruckler (1993) and compared with measured data to explore the effective depth for remote sensing. The parameters of the soils about Clay, Loam and Silt clay loam are provided but the considered depths are relatively thick and the parameters have a piecewise form depending on water saturations. Therefore, soil parameters are averaged and rounded. For the simulations in this study [Figure 3.9], n = 1.2 1.7 1.35, θ_s = 0.3 for all, T_s=35 C°, T_a=35 C°, RH = 0.4 and wind speed = 1.5 m/s. The parameter n for clay is reported as 1.5 but it leads a pore size too big and the most granulometric tables indicate that it can be smaller than silt-clay-loam. Therefore, we used 1.2 for this simulation. The case of Lehmann et al. (2018) and Merlin et al. (2016) provides various data sets with diverse parameters. For this study, several soil groups are selected among the sets, to test this model (resistance scheme) [Figure 3.10]. We need to note that the simulation scheme in Lehmann et al. (2018) is similar to our total resistance model except the soil pore heterogeneity. The major difference can be the usage of a equation

($\frac{\partial h}{\partial z} = 1 + \frac{e_0}{4K(h_c)}$) which is related to the characteristic length introduced by Lehmann (2008) and which is additionally applied from Eq. (3.16) to Eq. (3.18) (Lehmann et al., 2018).

3.5 Results and Discussion

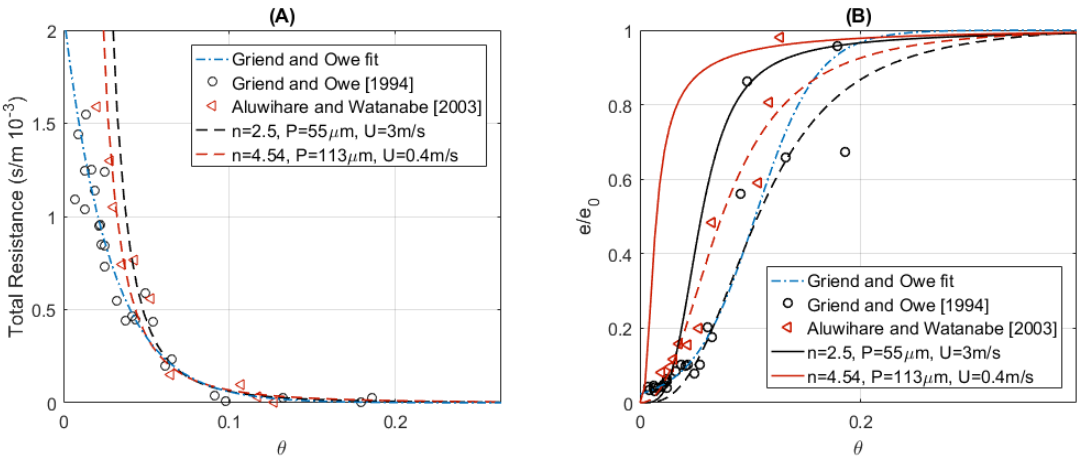

Figure 3.7: Plots in (A) are total resistance ($r_t = r_s + r_{BL}$) using Eq. (3.14). (B) is relative evaporation from Eq. (3.17) or Eq. (3.18). In plot (B), lines are simuated based on original SCH model Eq. (3.4) and dotted lines are from modified SCH Eq. (3.13) which is applied effective pore sizes and effective water contents. Marks are measured data by Aluwihare and Watanabe (2003) and Griend and Owe (1994). Additionally, fitted line by Griend and Owe (1994) also presented on the graph.

The total resistance models ($r_t = r_s + r_{BL}$) Eq. (3.14) for both SCH or modified-SCH (MSCH) tend to reasonably follow actual behaviors of the evaporation, and they demonstrate that the models have ability to cover a wide range of soil group (Haghighi et al., 2013). A plot using resistance terms cannot distinctly display evaporation rates because the resistance dramatically increases at some point by the action of low hydraulic conductivity. Therefore, the models also additionally simulate the relative evaporation rate [Figure 3.7; Figure 3.8; Figure 3.9; Figure 3.10]. In the all comparisons, the relative evaporations e/e_0 are estimated through the final form Eq. (3.17, 3.18) based on two different equations: original SCH Eq. (3.4) (line) and MSCH Eq. (3.13) (dotted line).

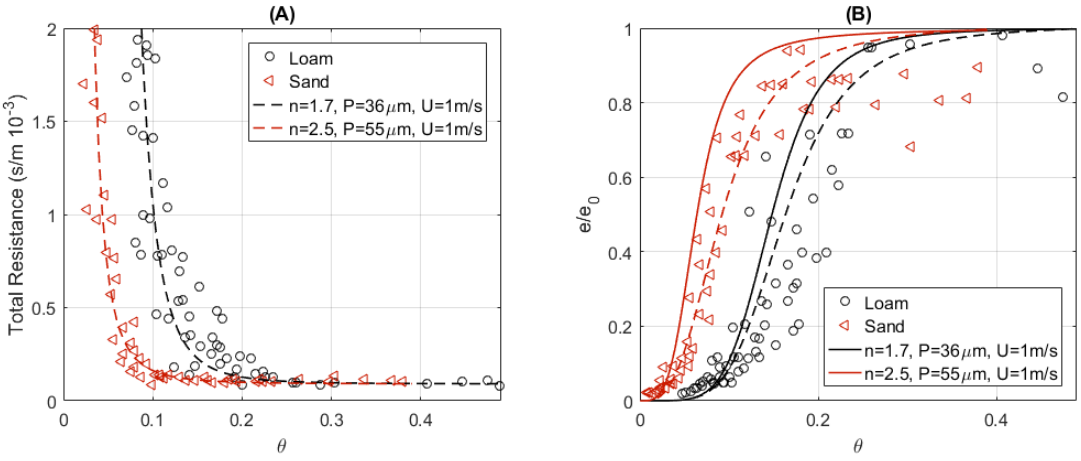

Figure 3.8: Plots in (A) are total resistance using Eq. (3.14). (B) is relative evaporation from Eq. (3.17) or Eq. (3.18). In plot (B), lines are simuated based on SCH model Eq. (3.4) and dotted lines are from modified SCH Eq. (3.13) which is applied effective pore sizes and effective water contents. Marks are measured data by Kondo et al. (1990).

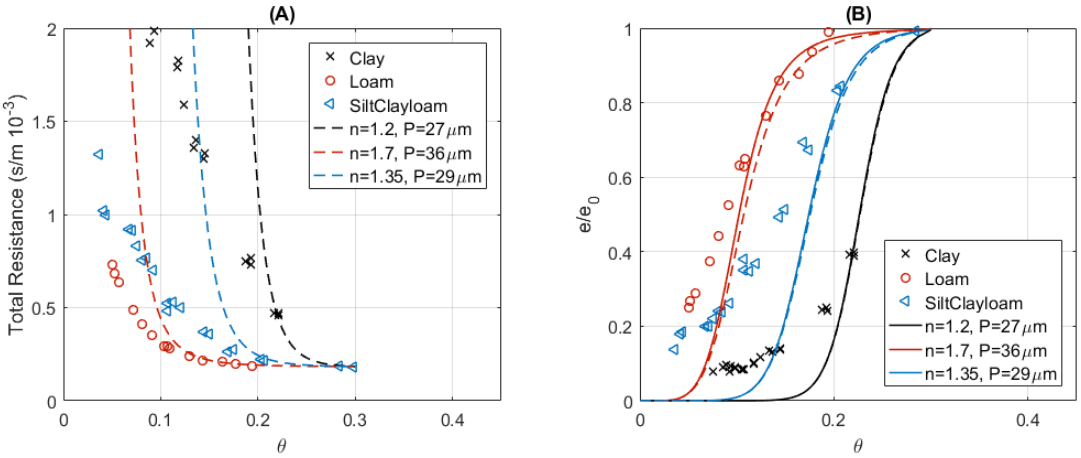

Figure 3.9: a similar graph as [Figure 3.7] and [Figure 3.8]. Marks are measured data by Chanzy and Bruckler (1993).

The scaled evaporation (dotted line on the graph (B)) through inhomogeneous soil tends to have smaller evaporation rates with observations than the results from the homogeneous soil (line). The important point in this test is that applying the heterogeneous soil pores can significantly affect the Stage I evaporation process. The two different simulations (line and dotted line) are only close to each other when soil has low hydraulic conductivity (K_s): soil resistance is dominant.

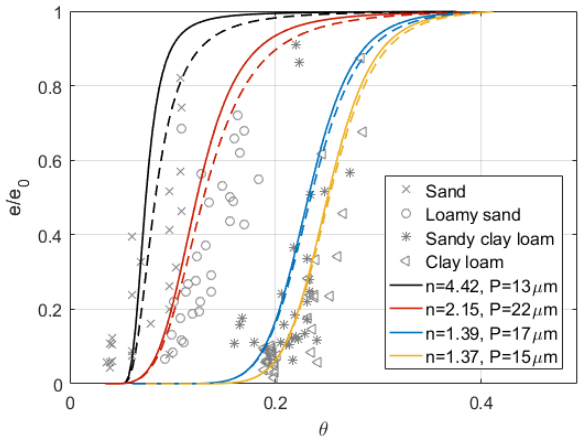

Figure 3.10: lines are simulated based on original SCH model Eq. (3.4) and dotted lines are from modified SCH Eq. (3.13) which is applied effective pore sizes and effective water contents. Additionally, $\frac{\partial h}{\partial z} = 1 + \frac{e_0}{4K(h_c)}$ is applied following Lehmann et al. (2018). The datasets and parameters are produced by Lehmann et al. (2018); Merlin et al. (2016).

The heterogeneous pore size and wet area yield different outputs in our simulation [Figure 3.11]. Therefore, applying the proper features of soil surface such as pore size, wet area, and other soil parameters can be important for the performance of evaporation models. This change in the model affects the evaporation rate in the middle of transitional stage (between stage I&II) but we can also see that the impact starts at the very beginning of evaporation process (nearly saturated) [Figure 3.7; Figure 3.8; Figure 3.11]. Such a range cannot be ignored because it is related to a large amount of water movement and it can lead a time dependent simulation being unsuccessful by inaccurate prediction of soil water contents.

The limiting supply of water from the soil r_s is still an major mechanism for change in evaporation stage and evaporation rate [Figure 3.11]. Under unsaturated soil condition, during an role of the boundary layer (r_{BL}) fades away, the limit of water supply in soil system by hydraulic conductivity $K(\theta_{surf})$ becomes the more dominant process. On the other side, the remaining evaporation activity after water supply is completely disconnected $K(\theta_{surf} = 0)$ cannot be captured yet [Figure 3.7; Figure 3.8; Figure 3.9; Figure 3.10]. This process may be caused by vapor transport inside of soil system based on front depths or evaporation depths so that air-invaded pore models

(Shahraeeni et al., 2012) can be applied to mimic such phenomenon.

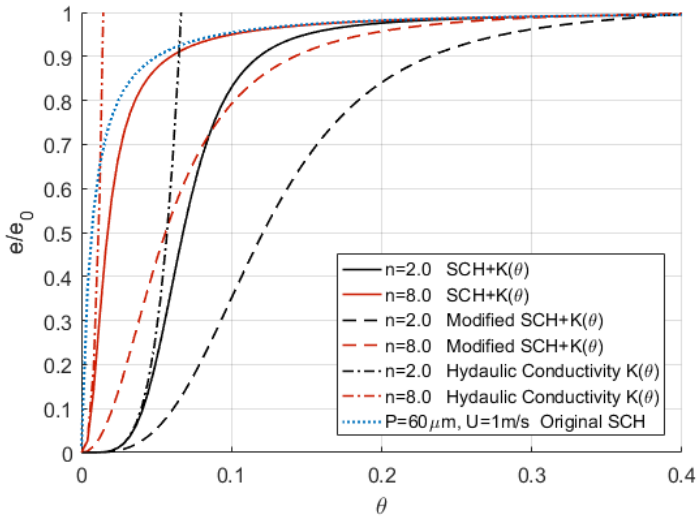

Figure 3.11: plots to see influence of effective water contents, effective pore sizes and hydraulic conductivity. Applying SCH means the resistance of boundary layer (r_{BL}) is active, and K(θ) (Hydaulic Conductivity) means soil resistance (r_s) is applied. Modified SCH means applying effective pore size and effective water content.

These all results indicate that the both behaviors r_{BL} and r_s can coexist under unsaturated conditions (Haghighi et al., 2013), and that, depending on soil properties and atmospheric condition, either resistance (r_{BL} or r_s) can dominant the process [Figure 3.7; Figure 3.8; Figure 3.9; Figure 3.10]. With small n and with small pore sizes, the evaporation rate can be mainly controlled by r_s due to the relatively low hydraulic conductivity compare with the atmospheric demand r_{BL}. These effects can be the cause of very short stage-I. With large n, which represents homogeneous pore size distribution and usually consists of large pore sizes, the boundary layer resistance r_{BL} tends to control the simulation, and the effect of heterogeneous soil grain can appear at this point [Figure 3.10; Figure 3.11]. This is because resistance of the water supply under the surface is not significant by large K, even though r_s would always be dominant at the end (very dry condition).

3.6 Summary and Conclusion

In this study, the relative evaporation models are explored and those modifications are introduced to reflect heterogeneous pore sizes of soil via capillary theory and pore-size distribution for an advanced evaporation model. The models we consider here are based on a theory that drying rate is controlled by resistance inside the soil r_s and resistance between the surface and the atmosphere r_{BL} as Haghighi et al. (2013). Above the soil surface (VBL), S&M model and SCH model are often used and they are regarded as having similar behaviors but different outcomes. The two models are compared and their similarities were explored. Moreover, many studies assume soil grains are homogeneous at the surface interface in VBL models, so pore sizes are regarded as a constant during their simulations of evaporation. To consider heterogeneous pore in this study, the scheme of different pore sizes are applied based on Young-Laplace equation and pore size distributions. In the same way, the 2D surface water contents are formulated corresponding to the different pore sizes.

Through the investigation, we found that the SCH model is the more practical method than S&M model for homogeneous soil grains. To compare SCH with S&M, the two models are simulated with the original concept [Figure 3.1]: (1) SCH is based on a fixed mean pore size P and a varying total pore size W, and (2) S&M is based on a fixed total pore size W and a changing active pore size P. If soil pores have the same size (homogeneous), under any state of water content, active pore size should not be changed and the number of inactive pores increases over the evaporative surface. Therefore, we may conclude that the considered area of horizontal soil surface might be expanded by increasing inactive pores like as SCH rather than decreasing active pore size as in S&M, while the water content decreases. Because this idea accords with SCH model, S&M model is modified to link the two models. The comparison through Eq. (3.4) and Eq. (3.5) shows that they are well matched [Figure 3.3]. The S&M model is mathematically derived based on Fick's law but a numerical method is required for the calculation. On the other hand, the SCH model is an approximate solution, so a simulation through SCH is straightforward and gives a reasonable performance compared to S&M. The investigation indicates that employing SCH method can be a

practical way to simulate evaporation rate and to reduce computational time. If one may want to use S&M model, a modification like Eq. (3.5) is recommended for such homogeneous case.

Second, inhomogeneous soil grains are applied on the SCH model, which results in the change of relative evaporation curves, especially at Stage I and its transition point heading to Stage II. To apply heterogeneous pores, effective pore sizes P_{eff} (the mean active pore size) are applied inside SCH model. As in [Figure 3.5], applying this idea tends to reduce the mean pore sizes and it can lead the relative evaporation rate increase according to the equation of SCH model (reduction rate decreases while the surface is drying). In a case with large n or homogeneous materials, the models can be close to the original SCH model. Effective water contents θ_{eff} are also calculated based on the derivation of the effective pore sizes and applied into the SCH model for heterogeneous soil evaporation. Considering both effective pore sizes and water contents with SCH model, the relative evaporation rates tend to be lower than original SCH model [Figure 3.6]. Although shrunk pore sizes by P_{eff} in the model should yield reduction of the decreasing rate of evaporation from Stage I to II, the simulated results show opposite behavior after the application of effective water contents. Therefore, effective water content, which is related to contact area between wetted surface and the atmosphere, can have more significant role in the mechanism of evaporation than the effective pore size.

The simulation is tested based on resistance models implemented by Haghighi et al. (2013) and Lehmann et al. (2018), and the results well follow measured data from several studies [Figure 3.7; Figure 3.8; Figure 3.9; Figure 3.10]. Except when the soil water content is close to zero, this study shows the resistance scheme is a reasonable method to capture both VBL and SL roles for evaporation, and it demonstrates that the models may have ability to cover a wide range of soil groups. The limiting supply of water from the soil r_s is an important mechanism. However, before the water supply for the evaporation is limited by low $K(\theta_{surf})$, the drying rate mainly depends on the process within the boundary layer, which controls large amount of water movement with even small changes. Through this test, in this period (Stage I&II), the inhomogeneity of soil pores can have a substantial impact on the evaporation rate.

Through this study, existing evaporation models are explored and modified, to find a practical diffusion model and to check its different behavior by applying heterogeneous pore sizes with various soil types. The relative evaporation model in this study shows reasonable outcomes compared with measured data. However, model error still exists: evaporation still occurs in actual data even when the hydraulic conductivity is zero in the models. Such behavior cannot be exactly explained yet, whether it is due to the diffusion process inside soil or it is still hydraulically connected, and how such processes actually work. For the future research, the models may need to include the influence of vapor fluxes within soil and temperature gradients, which may be related to a transition from Stage II to Stage III while evaporation process. Also, more experimental data about diverse soil types are needed with detailed information such as meteorological variables and soil parameters, to improve and verify evaporation models.

4. MODELING LAND SURFACE PROCESSES OVER A MOUNTAINOUS RAINFOREST IN COSTA RICA USING CLM4.5 AND CLM5

4.1 Introduction

Tropical forests play a critical role in determining regional and global climate. Due to their significance for the global water (Zhang et al., 2010; Choudhury and DiGirolamo, 1998) and climate cycles (Huntingford et al., 2013; Beer et al., 2010), accurate modeling of tropical regions is important for the prediction of future climate and climate change impacts. Forests in the tropics house 25% of the carbon stocks found in the terrestrial biosphere, accounting for 33% of global net primary production (NPP) (Bonan, 2008). While tropical forests occupy only 16% of the global area, they account for 33% of terrestrial evapotranspiration (ET), of which 70% is transpiration (TR) (Schlesinger and Jasechko, 2014). In tropical forests, ET can range from 1,000 mm up to 2,200 mm per year (Kume et al., 2011; Fisher et al., 2009; Loescher et al., 2005; Sheil, 2018). Hydrological processes in the humid tropics are also distinctly characterized by warm, uniform temperatures, large inter-annual and spatial variability, intense rainfall, and greater energy exchange accelerated by low albedos and high evaporative cooling (Wohl et al., 2012; Bonan, 2008). The loss of such forests by climate change or human impact can therefore be influential not only locally, but also in more remote regions (Lawrence and Vandecar, 2014). The accurate prediction and assessment of these extratropical effects and subsequent feedback to the reainforest are beyond the scope of this paper.

Land-surface models (LSMs), as a component of Earth system models (ESM), simulate the emission of heat, water vapor, and carbon dioxide from the terrestrial surface to the atmosphere, based essentially on the partitioning of net radiation (Wang et al., 2016). As an example of their application, they have been used for the assessment of the impacts of future climate change on tropical and extra-tropical forests (Cox et al., 2013; Huntingford et al., 2013). However, such models do not yet successfully capture the underlying complexity of land-atmosphere interactions

(Cai et al., 2014; Wang et al., 2014; Lawrence et al., 2011; Oleson et al., 2010). In particular, LSMs are known to make significant errors in the prediction of carbon and water fluxes for tropical regions, the reasons for which are not entirely clear, even though significant improvements have been made in this field of study (i.e., empirically and mechanistically). Lawrence et al. (2011) compared estimates obtained using two versions of the Community Land Model (CLM3.5 (Oleson et al., 2008) and CLM4.0 (Oleson et al., 2010)) to observed sensible and latent heat flux data from FLUXNET (Baldocchi et al., 2001). They found that CLM4.0 improved predictions compared to CLM3.5 for most sites across the network, but continued to show low agreement for tropical sites. Bonan et al. (2011) updated CLM4.0 by modifying the structure of radiative transfer model and physiological parameters for canopy processes, which resulted in notable improvements in CLM4.5 ((Oleson et al., 2013)) but overestimation of carbon/vapor fluxes persisted in areas closest to the equator. The deficit is especially true for tropical wet mountain rainforests, which have rarely been studied in the context of improving global LSMs, due to the lack of long-term/uniformly distributed measurement and the small number of observation sites (Fisher et al., 2009; Wohl et al., 2012).

To improve land surface models addressing tropical ecosystem biosphere-atmosphere interactions, partitioning net radiation (energy) and water accurately is critical for these models, especially with respect to estimating latent heat flux. Many studies maintain that vapor fluxes in the tropical site are highly correlated ($\approx$ 87%) with net radiation (Andrews et al., 2019; Fisher et al., 2009; Hasler and Avissar, 2007; Loescher et al., 2005). Others found that leaf wetness is also an important control (Andrews et al., 2019; Giambelluca et al., 2009). Some studies indicate that the effects of leaf wetness (which can contribute 8%-20% of ET) can appear depending on canopy water storage capacity and rainfall pattern, although short duration and high intensity rainfall does not significantly affect canopy evaporation (Kume et al., 2011; Loescher et al., 2005). In other words, water/vapor (e.g., ET/interception) related models are essential for tropical sites. Aerodynamic conductance has also been considered as a strong driver for evapotranspiration in tropical forest because the large amount of precipitation and frequently wetted canopy conditions control

leaf conductance (Shuttleworth, 1988; Loescher et al., 2005). Vapor pressure deficit (VPD) has been shown to only slightly influence ($\approx$ 14% predictor) on tropical ET (Fisher et al., 2009; Kume et al., 2011). However, when assessing these studies, it is noticeable that they all highlight the importance and difficulties of quantifying canopy water fluxes. ET dynamics are dependent on how these micrometeorological variables are related to the latent heat flux within the energy balance. In tropical forests, the Bowen ratio is consistently less than one (Loescher et al., 2005), which implies net radiation highly correlated with latent heat flux. Moreover, the forest canopy acts like a well-watered crop without water limits (Loescher et al., 2005; Hasler and Avissar, 2007; Kume et al., 2011). Tracking water movement within the system (water balance), finding ET proportion of net radiation (energy balance), and the deeper partitioning of the ET are still a critical question in highly saturated systems.

Therefore, accurately capturing each detailed hydrological processes, including partitioning rainfall flux by the sub-models (e.g., interception model or infiltration model), is a preemptive process for accurate tropical LSMs, not only because ET from the water has great portion of total energy exchange (Net Radiation) in fine scale, but also ET is a substantial portion of the water balance. Tropical regions, particularly in montane forest, play a crucial role in regulating and gathering atmospheric moisture in larger sense (Wohl et al., 2012). Clark et al. (2015) explores 11 LSMs and suggests possible improvement about detailed components of hydrologic processes, such as for soil moisture and ground water dynamics and including the space-time variability for plant activities. Also, the problem of spatial heterogeneity and hydrologic connectivity is mentioned, between large scale process (land-atmosphere flux) and microscale process (biophysical-biogeochemical-hydrologic processes) (Clark et al., 2015).

To the contrary, water is not our only concern, and it cannot be independently considered in Earth system or land surface system models. Other energy-related/physiological elements, such as thermal flux, radiative transfers, and vegetation activity (e.g., photosynthesis or respiration) are likewise important, because they are mutually dependent on the water. The main schemes of any land model including CLM (e.g., heat/vapor flux, carbon flux, and net radiation) and additional

components have been continuously incorporated and updated in LSMs to represent more realistic processes and new community interests, including: soil carbon and nitrogen cycling (Thornton et al., 2007), multi-layer plant canopies (Ryder et al., 2016; Launiainen et al., 2015; Bonan et al., 2018), and even more sophisticated urban/ rural model (Lawrence et al., 2018; Buzan et al., 2015). These indicate that the current LSM consists of a plethora of sub-models, making it difficult to identify a specific sub-model or set of sub-models leading towards an entire model error.

Hence, in order to properly parametrize global LSMs and to precisely represent such complicated systems, such as the tropics, it is necessary to keep diagnosing the land surface models relying on site-based data especially from tropical sites, and such site-based studies have confirmed in developing advanced LSMs (Mao et al., 2016). Unique site like tropical forests are valuable testbeds for these models because the environment is beyond the normal boundary condition, which enhances model calibration under more extreme climate conditions associated with climate change. In site-based studies via data analysis (e.g., simple comparison plots or diurnal variation plots), it is possible to identify and alleviate the error of sub-component model (such as soil moisture/temperature, interception, stomatal process, etc.). Such errors cannot be easily detected by the analysis of main/large-scale elements (such as albedo, net radiation, etc.).

Land surface models are gradually increasing their resolution with the improvement of remote-sensing technology in which reflects the high heterogeneity of the Earth surface. This exacerbates current model uncertainty caused by blurred/oversimplified parameterization and misinterpretation (Singh et al., 2015; Wood et al., 2011). Site-based measurements including eddy-covariance tower systems have been widely utilized for the advance of global land surface models via calibration and validation (Bonan et al., 2012; Zaehle and Friend, 2010; Larsen et al., 2016; Chaney et al., 2016), and such observations will be continuously needed. Gridded global data from the FLUXNET network is also available for model development at large scales (Bonan et al., 2011; Jung, 2009). However, point-scale or small scale study is still the core and the base-line for any other scale study. In this study, CLM4.5 (Oleson et al., 2013) and its updated version (CLM5) (Lawrence et al., 2018) are employed, and micrometeorological datasets from a tropical rainforest in Costa

Rica are compared with these simulation results. The objectives are three-fold:

1. To compare the default, point-scale predictions of CLM 4.5 and CLM 5.0 against micrometeorological and flux measurements collected in a Costa Rican wet montane tropical forest;

2. To identify the improvements in performance between the two CLM versions, and to detect the still existing shortcomings in the newer version (CLM5), which errors are possibly caused by the unique environment at our study site (i.e., frequent rainfall and mountainous topography) and by simplified formulations and coarsely assigned parameters (i.e., interception/leaf-wetness models, photosynthesis models, etc.); and

3. To determine which canopy-atmosphere processes are most poorly represented, and identify which sub-models are responsible for the largest errors in radiation partitioning, in order to suggest priorities for future model improvements.

4.2 Methodology

4.2.1 Study Site

The field site is located at the Texas A&M University Soltis Center near San Isidro de Peñas Blancas in Costa Rica ($10°23'13''N, 84°37'33''W$, about 600 m above sea level). This area has a mean annual temperature of 24 °C, relative humidity of 85%, and precipitation of 4200 mm (Teale et al., 2014). The study area is classified as a transitional, tropical premontane, moist forest. The canopy height ranges from between 24 and 45 m, and its terrain has a steeply eastern slope (Aparecido et al., 2016; Jung, 2009). Rainfall is frequent, and a little over two-thirds of days have one or more rain events.

4.2.2 Micrometeorological Measurements

The site has two primary biometeorological measurement locations. The main weather tower (hereafter called "Met Tower") is located in a flat clearing over grass at the base of the mountain. The walkup canopy access tower (hereafter called "Canopy Tower") is located within the forest, on the eastern slope. The Met Tower allows for the measurement of meteorological conditions

without the influence of canopy processes and resistances. Precipitation (mm; TE525, Campbell Scientific, Logan, UT), incoming solar radiation, net radiation ($W \cdot m^{-2}$; CNR1, Campbell Scientific), air temperature (°C; HMP60, Campbell Scientific), and relative humidity data (%; HMP60, Campbell Scientific) have been collected since 2010. The Canopy Tower has collected the same variables as the Met Tower (with exception of precipitation) and additional measurements, such as vapor and CO_2 concentration, soil moisture, leaf wetness and sap flow since 2014. An infrared trace gas profiler (IRGA; AP200, Campbell Scientific) and an eddy-covariance system (LI-7200, LI-COR, Lincoln, NE; CSAT3, Campbell Scientific, Logan, UT) are used to collect micrometeorological data at various heights, including concentrations of water (i.e., H_2O) and carbon dioxide (i.e., CO_2), wind speed and direction, and air temperature. Additional data are also collected to track canopy processes: leaf wetness sensors at four different heights (LWS, Decagon Devices, Utah), photosynthetically active radiation (PAR) profiles (LI-190, LI-COR) at five heights, leaf area index (LAI) profile using a lined PAR sensor (LI-191, LI-COR) and Beer-Lamber law (Vose et al., 1995; Andrews et al., 2019), leaf temperature sensors for sunlit and shade leaves (SI-111, Apogee Instruments, Logan, UT), soil heat flux (HFT3, Campbell Scientific), soil temperature (5TE, Decagon Devices, WA), soil moisture (EC-4 and 10HS, Decagon Devices, WA), soil respiration (LI-8100A, LI-COR) and transpiration from sap flow system. Refer to Aparecido et al. (2016) and Andrews et al. (2019) for more detailed information about the sap flow system and the profile measurements, respectively. The datasets for this site from 2014 to 2017 are available online (Miller et al., 2018a,b,c,d).

The Canopy Tower is situated above the canopy but near an emergent tree [Figure 4.1], leading to a large gap in the canopy in between heights of roughly 30 and 40 m. This configuration leads to two main challenges. Below this gap, the canopy is closed. Above the gap, one emergent, upslope tree provides a significant degree of shading, which leads to a 70% reduction in PAR between measurements at the canopy surface (32 m) and above the emergent tree (44 m). We also note that this configuration has some implications which make the eddy covariance method less ideal. The sonic anemometer and IRGA are located at 34 m height, extending away from

the tower and clear of obstructions in both the upwind and downslope directions [Figure 4.1]. Predominant 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 ecosystem level variables, like gross primary production, they are suitable for testing and validating models despite the presence of such significant canopy gaps by the emergent tree.

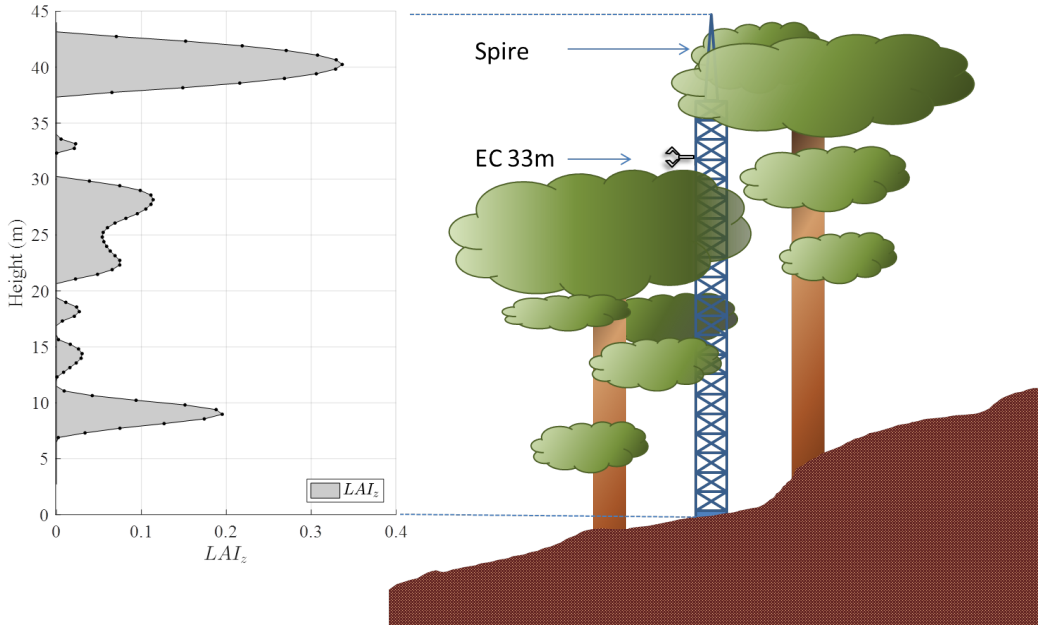

Figure 4.1: Sketch of Canopy Tower located in a plot within a mature premontane moist tropical forest in Costa Rica (right) with LAI profiles highlighted (left) along with the location of the eddy covariance system (EC 33m) and the spire (44 m) hoisting the net radiometer including an incoming solar radiation sensor. The sum of points (LAI_z) totals $LAI(\mathrm{m}^2 \cdot \mathrm{m}^{-2}) = 6$ for this stand. The method used to obtain the LAI profile of the canopy is explained in a previous study Andrews et al. (2019)

4.2.3 Model Description

In this section, the land model is briefly described focusing on the model structure and main formula (i.e., energy balance equation), and several sub-models which have high potential of causing

model errors for this site are examined. Considering that our study site has extremely high humidity and annual precipitation, the sub-models related to water fluxes are hypothesized as being the main sources of prediction/estimation errors. There are many sub-models within CLM, and more detailed descriptions can be found in the technical manual (Lawrence et al., 2018; Oleson et al., 2013, 2010). Additionally, hereafter we use CLM in a general sense, with a statement being applicable to both CLM4.5 and CLM5; but provide the specific version number when distinguishing their respective behavior or the effects of recent code modifications.

CLM calculates the radiative transfer through the canopy and the surface, using the Two-stream approximation method (Dickinson, 1983; Sellers et al., 1992; Bonan, 1996; Oleson et al., 2013), which is a starting point for land surface models determining the amount of energy exchange. In the procedure, the canopy structure and the albedo, as influenced by current conditions (leaf angle, wetness, solar angle, etc.), are main controllers determining the absorptivity of incoming solar radiation by the canopy and the soil surface. Based on the absorbed energy of incoming energy, fluxes of sensible heat, latent heat, and soil heat are estimated using the energy balance equation. For example, as a function of vegetation temperature (Tv), the canopy energy balance can be written as

$$-S_v + L_v(T_v) + H_v(T_v) + LE_v(T_v) = 0 \tag{4.1}$$

where Sv is the absorbed solar radiation by canopy, L_v is the long wave radiation emitted by canopy, Hv is the sensible heat flux, and LEv is the latent heat flux from the canopy, all of which are given in $W \cdot m^{-2}$ (Oleson et al., 2013). Monin-Obukhov Similarity Theory (MOST) is used to determine resistances along the soil-plant-atmosphere continuum [Figure 4.2], which is then used to calculate Hv and LE_v (Zeng et al., 1998; Oleson et al., 2013). As a big-leaf model, CLM represents both sunlit and shade leaves (Dai et al., 2004).

The water balance equation tracks the water movement through the system and connects to the energy balance via its dual controls on ET. The first of these controls, the influence of soil moisture on stomatal conductance, is not considered in this study. Prior work has determined that ET at the

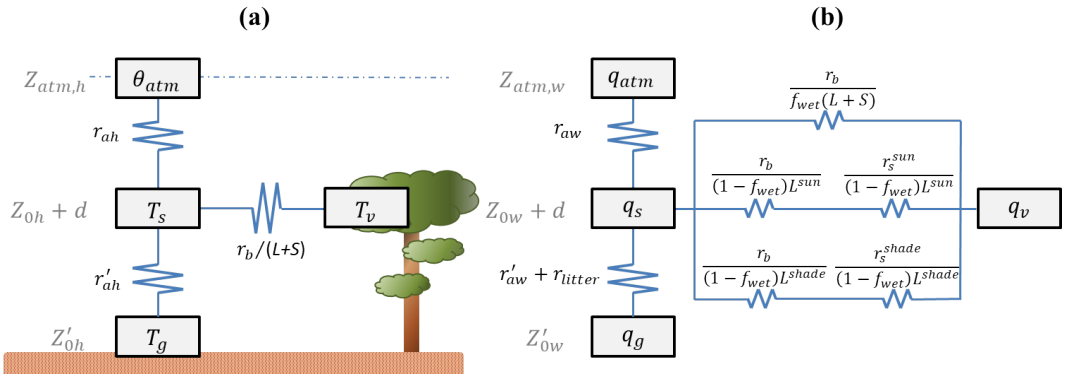

Figure 4.2: Resistance network schemes incorporated within CLM for a) sensible heat flux and b) latent heat flux. Main state variables are atmospheric forcing temperature (θ_{atm}) and specific humidity (q_{atm}), canopy air temperature (n) and specific humidity (q_s), leaf temperature (T_v) and its corresponding specific humidity (q_v), and ground temperature (Tg) and its corresponding specific humidity (q_g). Relevant heights are the atmospheric reference height (z_{atm}), the canopy roughness heigth (Z_0), the groundwater roughness height (Z'_0), and the displacement height (d). Resistances are specified by their scalar (h for heat and w for water vapor), type (a for aerodynamic, b for boundary layer, s for stomatal, or $litter$ for litter), and lighting (sun or $shade$). Leaf wetness also exerts control on fluxes, via a wetness fraction (f_{wet}) and ($L+S$) is leaf and stem are index. Figure adapted after Oleson et al. (2013).

present study site is not limited by soil water deficits during normal to above-normal rainfall years, such as the period from 2014 to 2016 (Andrews et al., 2019). Rather, while its effect is considered to be small in some ecosystems (Burns et al., 2018), previous studies have shown that leaf wetness exerts significant influence on fluxes from rainforests in general (Loescher et al., 2005; Kume et al., 2011) and in this site specifically (Aparecido et al., 2017; Moore et al., 2018). CLM reflects these mechanisms as well in the resistance network [Figure 4.2b], and the leaf wetness prevents transpiration and contributes to canopy evaporation rates. Here, leaf wetness is determined by the interception rate of incoming precipitation (Deardorff, 1978; Dickinson et al., 1993; Lawrence and Chase, 2007). The amount of interception q_{ic} is given in CLM4.5 as:

$$q_{ic} = 0.25 \cdot q_{rain/snow} \cdot [1 - e^{-0.5(L+S)}] \tag{4.2}$$

and in CLM5 as:

$$q_{ic} = 1.00 \cdot q_{rain/snow} \cdot tanh(L+S) \tag{4.3}$$

where $q_{rain/snow}$ is the precipitation as liquid or snow, and 0.25 is a model coefficient. Here, we note that when the leaf-stem area index is high ($L + S > 2$) the interception rate approaches 100% in CLM5 only. This can be questionable in our view because canopy in this site, which has LAI far higher than 2 ($m^2 \cdot m^{-2}$), does not cover 100% sky ($\approx tanh(2)$). On the other hand, the value of 0.25 in CLM4.5 seems too low. After determining intercepted rainfall, canopy water storage (W_{can}) is calculated through re-partitioning based on the condition of $0 \leqq W_{can} \leqq W_{max}$, where W_{max} is $0.1(L + S)$ (Dickinson et al., 1993; Oleson et al., 2013). Finally, f_{wet} is

$$f_{wet} = \left[\frac{W_{can}}{W_{max}}\right]^{2/3} \tag{4.4}$$

Additionally, in CLM5, f_{wet} cannot exceed a maximum value (f_{wetmax}) of 0.05, and f_{dry} is calculated as:

$$f_{dry} = \frac{(1 - f_{wet}) \cdot L}{L + S} \tag{4.5}$$

In Eq. (4.4), the 2/3 exponent was assumed following the original literature (Deardorff, 1978), because the canopy water tends not to be evaporated when it is set to one and evaporates too fast when set to zero (Deardorff, 1978).

Additionally, CLM mainly uses the Farquhar model (Farquhar et al., 1980; Oleson et al., 2013) for the estimation of photosynthesic rates. In our site, air temperature varies little throughout the year, and CO_2 concentration is not significantly variable. Consequently, light-limited photosynthesis can be considered as a dominant process. The light-limited model w_j ($\mu mol \cdot m^{-2} s^{-1}$) in CLM is developed based on the Farquhar model (Oleson et al., 2013) and can be written as:

$$w_j = 0.25 J_x \mathbb{C}_i, \qquad \mathbb{C}_i = \frac{c_i - c_p}{c_i + 2c_p} \tag{4.6}$$

where c_i is intracelular CO_2 concentration, c_p is CO_2 compensation point, assuming 4 electrons per CO_2 molecule, $\mathbb{C}_i$ is a function of c_i and c_p, and J_x ($\mu mol \cdot m^{-2} s^{-1}$) is the electron transport

rate which can be additionally estimated through

$$\Theta J_x^2 - (I_{PSII} + J_{max})J_x + I_{PSII}J_{max} = 0 \tag{4.7}$$

where Θ is a curvature parameter ($\Theta = 0.7$ in default), J_{max} ($\mu mol \cdot m^{-2}s^{-1}$) is maximum rate of electron transport, I_{PSII} can be estimated as $I_{PSII} = 0.5\Phi \cdot I_{APAR}$, Φ is quantum efficiency of photosystem II ($\Phi = 0.85$), 0.5 is for two photosystems for one electron, and I_{APAR} is absorbed PAR ($\mu mol \cdot m^{-2}s^{-1}$).

To further explore these relationships, Eq. (4.6) and Eq. (4.7) are simplified and recalculated to make them comparable to apparent quantum yield (α). This is because the light-limited model has a hyperbolic shape caused by other environment conditions. However, the apparent quantum yield is a slope parameter (or the initial slope of the light-limited model) between absorbed-PAR and photosynthetic rate, which we can see it is a well known and simple parameter with a long research history in the literature (Skillman, 2007; Evans, 2013). From Eq. (4.6), if ambient condition has $c_p \approx 40\mu mol$, $c_a \approx 400$ and $c_i/c_a \approx 0.7$,which gives $c_i \approx 0.7 \cdot 400\mu mol$ (Launiainen et al., 2011; Katul et al., 2010), then $\mathbb{C}_i$ becomes 0.667. Also, if c_i becomes higher as atmospheric CO_2 concentrations increase, it becomes closer to 1. Through Eq. (4.6) and Eq. (4.7), the initial quantum yield of CO_2, also known as apparent quantum yield (α), can be estimated via $\partial J_x/\partial I_{APAR} \times 0.667 \times 0.25$, which can be used as simple-version models such as $w_j = \alpha \cdot I_{APAR}$. The theoretical maximum for α should be ≈ 0.11, α with saturated condition is approximately 0.075 (absence of photorespiration), and in normal atmosphere condition α is about 0.05 which is estimated if $\Phi \approx 0.6$ in Eq. (4.7) (Evans, 2013; Raj et al., 2015; Skillman, 2007). These light-limit models with different parameters are explored with observations in later section.

4.2.4 Simulation Setup and Comparison Method

CLM was tested in point-scale mode, using the satellite phenology (SP) with default settings, with exceptions noted below. Extension modes, which consider additional processes such as dynamic global vegetation (DGVM), biogeochemical cycles (BGC), or carbon-nitrogen cycling (CN)

were mostly not considered since they do not afect our study interests here (e.g., tree growth and stand competition). Input parameters for the simulation were determined using the 'mksurfdata_map ' utility provided in Community Earth System Model (CESM). The utility derives its values from satellite-based global datasets of phenology, soils, and topography, provided by University Corporation for Atmospheric Research (UCAR) (Oleson et al., 2013).

Based on many tests, we decided to use global parameters as possible if there was no significant influence on model performance. Location specific default parameters from the global dataset included: leaf area index (LAI, $5m^2 \cdot m^{-2}$), stem area index (SAI, $0.8m^2 \cdot m^{-2}$), canopy height (34 m), sand clay loam soil (47% sand, 26% clay, 27% silt), organic matter density (33 $kg \cdot m^{-3}$), maximum fractional saturated area (0.39), color class (15). We need to note that a small change of these parameters to apply local parameters did not affect much the model's results in our test. This is because our LAI value is high enough to be the dominant process, and the role of the soil is small. Moreover, the slope parameter exists but it is never used in the radiative transfer, canopy process, and turbulence model in CLM. Also, the most of the measured parameters at this site are not much different from the default values. Therefore, we decided to use the default setting except for some significant differences as below. The tropical, broadleaf evergreen tree (BET) plant functional type (PFT) was used as the basis for representing the site's specific landcover. The location in question had a default value of 30% BET tropical, 30% of tropical broadleaf deciduous trees (BDT Tropical), and 25% for grass and crop which we altered to 100% BET for purposes of this study. The atmospheric reference height was set to 44 m to reflect the location of the net radiation sensor on the Canopy Tower.

As an input, a meteorological forcing data set for CLM was created based on the measurements collected on site. These variables include half-hourly averages of wind speed ($m \cdot s^{-1}$), incoming solar radiation, relative humidity, air temperature, air pressure, precipitation, and CO_2 concentration. Comparison of the simulation was based on measurements taken at Canopy Tower; thus, Canopy Tower data was primarily used as forcing data when possible data was available. Average precipitation and air temperature data collected at 10-m height at Met Tower were also used for

data gap-filling. In most cases, weather data obtained from the two towers were highly correlated, as the locations are less than 1 km away and only differ in their immediate surroundings (i.e., forest vs. clearing) and slope degree (i.e., $\sim$45 degree slope vs. flat terrain).

Although flux methods cannot measure gross primary production (GPP) directly, it is an extremely important variable in the context of global carbon cycle modeling. In light of this, we estimated GPP based on net ecosystem exchange (NEE), net ecosystem production (NEP), and ecosystem respiration (ER), where NEE $\approx$ NEP and GPP = NEP – ER. With eddy-covariance data being collected at the height of 33 m, NEP was estimated as CO_2 flux + CO_2 storage flux. Ecosystem respiration (ER) was estimated to be around 1.2 ($\mu mol \cdot m^{-2}s^{-1}$) based on the nighttime data found using the u* threshold method (Papale, 2006; Reichstein et al., 2005). This EC based data for CO_2 flux including H_2O flux can be still questionable due to the circumstance of the measurement. However, later comparison shows these data provide acceptable tolerance when compared to sap-flow data, and they are enough to give the information whether the model is over-parameterized.

For transpiration (TR), measured data and simulated transpiration rates are compared at daily timescales. To investigate water loss from the canopy, it is necessary to estimate or measure each major flux (partitioned flux) within ET. In this site, up-scaled sap-flow data provides a transpiration rate (Aparecido et al., 2016), which in turns allows for water vapor flux partitioning. Although the sap-flow data at the site tends to be temporally lagged and a nocturnal sap-flow activity appears (shown later), it provides data to be used as a comparison at a daily scale against CLM estimates. As CLM cannot represent nighttime transpiration, sap flow measurements collected when the cosine zenith in CLM is less than zero were eliminated from the comparison. This daily scale comparison is made by a one to one figure with R-squared value, and also regression analysis which provides additional information how much the model is deviated from observation: the slope and its intercepts supposed to be 1 and 0. Particularly, the intercept is related to the mean value so it is directly affected by whether the nighttime transpiration is included in this analysis. This difference can reflect the lag between sap-flow at the base of the tree and actual transpiration rate.

In physiological model for GPP and TR, unlike the radiative transfer models, notable updates to the physiological models and their associated parameters have been made in CLM5 compared to previous version in CLM4.5. The Ball-Berry Model (BB) (Ball et al., 1987) was supplanted by a combination the Medlyn model for the stomatal conductivity (Medlyn et al., 2011), a plant hydraulic stress model (Bonan et al., 2014), and the Leaf Use of Nitrogen for Assimilation (LUNA) routine (Ali, 2016). While the BB model still can be used for CLM5, the slope parameter, which is directly associated with stomatal conductivity, has been changed from 9 to 7.3 for C3 plants. We have tested several options in CLM5 and determined that changing the stomatal conductivity model does not affect photosynthesis-related results (e.g., GPP) in our case. Therefore, we can simply understand that the slope of BB parameter has been reduced for CLM5 for carbon flux and vapor flux, if differences are identified in photosynthesis rate between two different CLM versions.

For some comparisons, the determination of each variable's height is necessary for CLM. In this case, each reference height is determined based on assigned parameters in CLM: the displacement height is $d = 23.45m$, ground roughness height is $z_{0mg} = z_{0vg} = z_{0hg} = 0.01m$, and surface height is $z_0 = z_{0mv} = 2.625m$, so the canopy height is $d + z_0 = 26.075m$. Therefore, for instance, canopy air temperature (Ta) in CLM is 2 m temperature in this comparison study, so it has $d + z_0 + 2 = 28.075m$. Our instrument heights do not exactly correspond to those assumed by CLM, so one or two closest-positioned data are used for the comparison rather than interpolating all data.

Additionally, in this simulation, since CLM5 has low leaf wetness ratio (the maximum is 0.05 as Eq. (4.4)), a range of leaf wetness parameters were applied to determine its influence. To explore leaf wetness models, all leaf wetness [0-1] from CLM were re-estimated based on the water amount on canopy using Eq. (4.4). Additionally, the question whether or not to apply the power of 2/3 does not change much our comparisons.

Soil related data is spatially up-scaled and vertically interpolated to compare with the simulation. For the spatial up-scale, soil temperatures and soil heat fluxes are measured at five different places near the Canopy Tower, and the vertical profile data are also collected close to the base of

the tower. For the vertical profile, CLM considers more number of soil layers, so the results of CLM are linearly interpolated, to compare with the measured data.

To initialize the simulations, CLM was first executed with a cold start (i.e., randomly produced initial values) and run for 100 years to get stable soil temperatures, cycling 6 years datasets collected between the beginning of 2010 and the end of 2015. Once stable soil temperatures were obtained, CLM was rerun for two years (2014 – 2016) at a 30-minute time step. Linear regressions were performed to compare CLM outputs to field data. Goodness-of-fit of the regression analysis was provided based on coefficient of determination (R-squared) if necessary. The variables considered when comparing observed vs. predicted were net radiation, PAR, albedo, CO_2 flux, GPP, transpiration, latent heat flux, air temperature, leaf temperature, leaf wetness, and soil related variables. Additional test simulations were used to determine how changes in levels of maximum leaf wetness (f_{wetmax}) and quantum efficiency of photosystem (Φ) affected goodness-of-fit. The modifications of LAI, light extinction related coefficients, and canopy heights (34m$\sim$44m) were also tested. Unlike f_{wetmax} and Φ, however, they provided no significant difference or better results, so comparison and discussion of them are not made here.

4.3 Simulation Results and Comparison

4.3.1 Net Radiation and Albedo

Light related comparisons indicate modeled land surfaces receive less energy than field measurements, even though net radiation components were overall well predicted. Net radiation simulations were -20 $W \cdot m^{-2}$ less than the average measured values, although diurnal patterns closely matched (R^2=0.99). CLM was approximately -15 to -45 $W \cdot m^{-2}$ lower than field measurements during the daytime, -10 to -15 $W \cdot m^{-2}$ during the nighttime [Figure 4.3a; Figure 4.3b], indicating that the land surface in CLM tends to absorb less energy than the actual surface. Little difference ($< 5\ W \cdot m^{-2}$, R^2=0.99) was detected between CLM4.5 and CLM5. The following simulated shortwave reflectance (albedo) in CLM is higher (+0.022 in all daytime data, around 15% higher) than the gauged albedo [Figure 4.3c], which can be also implied by the result of net radiation

(longwave and shortwave radiation) above.

Light data was clearly affected by the sloped terrain. Although the models were developed for all the global surface, sub-grid scale heterogeneity in land surface elevations has not yet been implemented in CLM4.5/5.0. Albedo from CLM tends to have a symmetric form, while the measured albedo has a skewed diurnal pattern [Figure 4.3c]. A large discrepancy occurs in the early morning and the largest during mid-afternoon (+0.0517 at 3PM; [Figure 4.3c]). Additionally, the reflectance in the morning is usually higher than in the afternoon [Figure 4.3c]. Except for the early morning and late afternoon, when solar angle or intensity peaks [Figure 4.3e], the albedo in CLM is close to the measured data. The highest PAR intensity (or highest incoming solar radiation) occurs at 10AM, at which time produces the smallest difference between observation and simulation of albedo (+0.0214 in [Figure 4.3c]). In some parts, this may be caused by the oversimplification of albedo models which cannot properly respond to the intensity of solar radiation/angle. However, the skewed albedo seen on the measured data but not in the simulation [Figure 4.3c; Figure 4.3d] clearly indicates CLM cannot represent the slope effect of the land surface. Such skewed diurnal variations are also observed in the PAR profiles [Figure 4.3e; Figure 4.3f]. The measured PAR values, which are somewhat shaded by the upper canopy, are diurnally skewed compared with shaded PAR in CLM. Different from the solar radiation above the canopy (i.e., the top of the tower), radiation starts to get skewed right after infiltrating the top canopy layer. If revisiting the effect of canopy gaps by the emergent tree for a quantitative check, we can check that radiation between the top of the canopy ($\approx$ 400 $W \cdot m^{-2}$ at 44 m from Net Radiation) and the next nearest heights ($\approx$ 110 $W \cdot m^{-2}$ at 32-38 m from PAR) are considerably different (about 70-80% reduction from the top) as mentioned before. The height of tall tree on this site is about 38 m (Aparecido et al., 2016), so the shade effect is significant within small thickness of canopy (the emergent tree).

4.3.2 CO_2 Flux (GPP)

All CLM versions (CLM4.5, CLM5, and CLM5BGC) overestimate GPP (6.7, 4.9, and 3.6 $\mu mol \cdot m^{-2}s^{-1}$) [Figure 4.4a; Figure 4.4b]. Compared to CLM4.5, results from the new version CLM5 seem to be more similar to measured data [Figure 4.4a]. However, CLM5 yields a lower

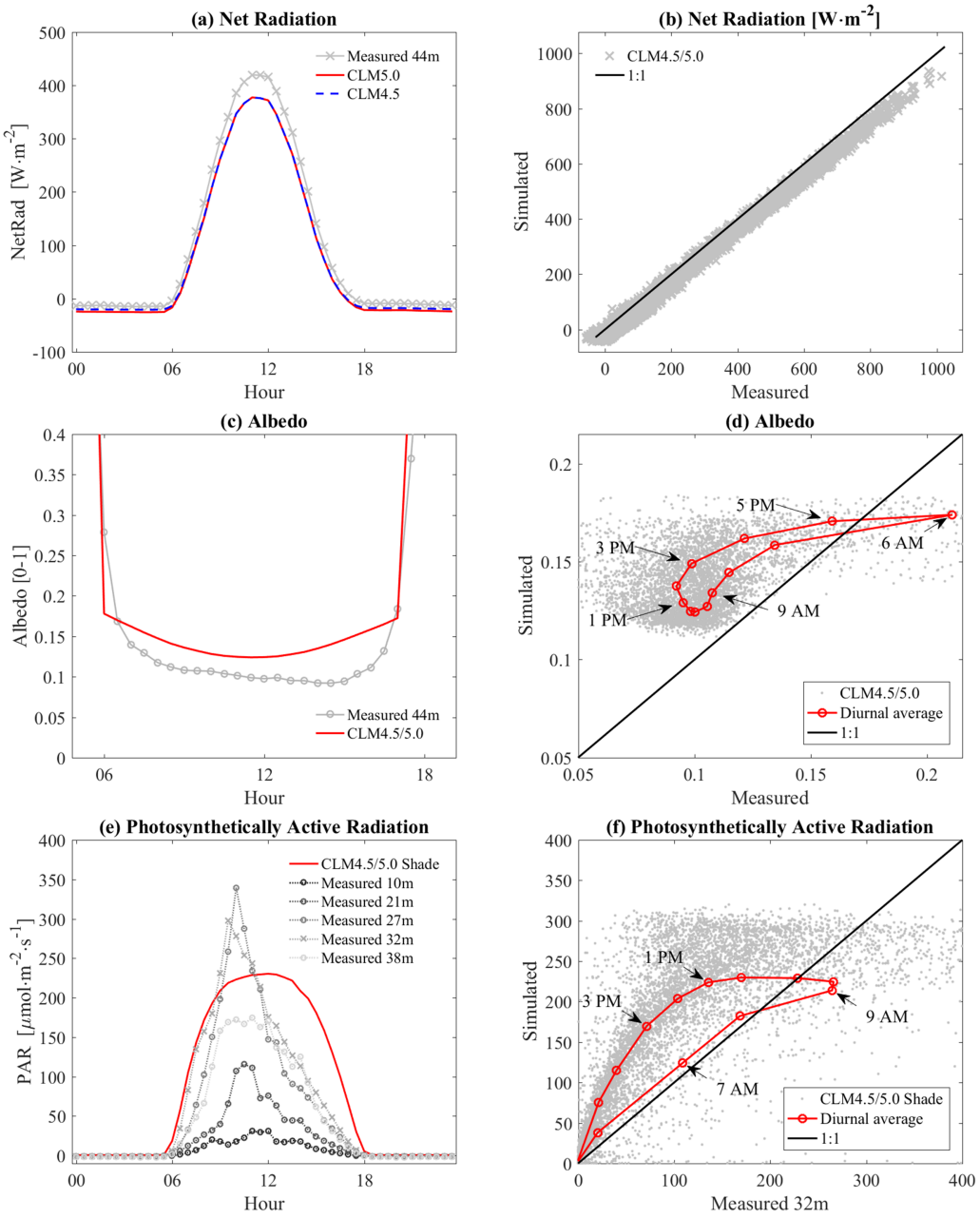

Figure 4.3: (a) and (b) - comparison of net radiation between CLM and measurement on Canopy Tower at 44m. (c) and (d) - albedo at 44m. (e) and (f) - PAR comparision for shaded canopies. All left plots (a, c, and e) are ensemble diurnal variation and the right plots (b, d, and f) are one to one comparison plots between CLM and measured data. Hysteresis depicted on (d) and (f) is based on hourly ensemble average values for daytime.

photosynthetic rates than CLM4.5, possibly due to the lower BB parameter and lower maximum rates of $V_{c,max25}$ and J_{max25} which the maximum rate is suppressed by LUNA and BGC mode.

Here in CLM5, disabling the plant hydraulic model increases the carbon-assimilation rate, while disabling the LUNA model decreases it in this site study. The prediction for the middle range of photosynthetic rate (5-15 $\mu mol \cdot m^{-2}s^{-1}$) has not been much improved compared with CLM4.5.

One of the possible causes of discrepancy between the GPP and its observation may be the light-limiting model [Figure 4.4b]. Comparison between absorbed PAR versus GPP shows that the initial slope of measured data is much lower than the simulated one [Figure 4.4c]. As previously explained (Figure 4.4d), through an extensive literature study Skillman (2007) and Evans (2013) show the theoretical maximum for α should be $\approx$0.11, α with saturated condition is approximately 0.075 (absence of photorespiration), and in normal atmosphere condition without any extreme condition α is about 0.05 which is estimated when considering Φ $\approx$0.6 in Eq. (4.7) (Evans, 2013; Raj et al., 2015; Skillman, 2007). Here, our fitted value for α is 0.021 (Φ $\approx$0.25) from our observation. This low value can be caused by other factors such as physiological stress or a scale problem. The fitted value is estimated from eddy-covariance measurement, not from leaf scale study. In current CLM4.5 or CLM5, the α is around 0.07 with $\mathbb{C}_i$=0.667, which is higher than 0.05 as usual cases (Skillman, 2007; Ehleringer and Pearcy, 1983; Ehleringer and Björkman, 1977). For this study, Φ was constrainedly modified to get proper α, but we need to revisit this model for the future study.

Test simulation with CLM4.5 and CLM5 was conducted using Φ=0.25 and Θ=0.7. By altering Φ value, both CLM4.5 and CLM5 had better performance compared to the previous comparison [Figure 4.5;Figure 4.4a]. The shape of the comparison plot in the middle range of GPP in this study (5-15 $\mu mol \cdot m^{-2}s^{-1}$) results in more stable prediction with new parameter. However, maximum GPP has been reduced as expected [Figure 4.4d], and it is possible to fix such reduction by updating Θ (curve shape), as shown on [Figure 4.4d]. In the simulated diurnal variation plot, the trend is slightly shifted in the afternoon, also probably due to the effects of the topographical slope [Figure 4.5b]. Time-dependent classification (i.e., regression lines with intercept forced through zero [Figure 4.5b]) and the fitted slopes indicate that geographical features have an influence on photosynthetic activity, which is mostly caused by the radiative transfer models, like albedo. However, the model fails to accurately represent such features, since CO_2 flux in CLM is lower in the

morning and higher in the afternoon.

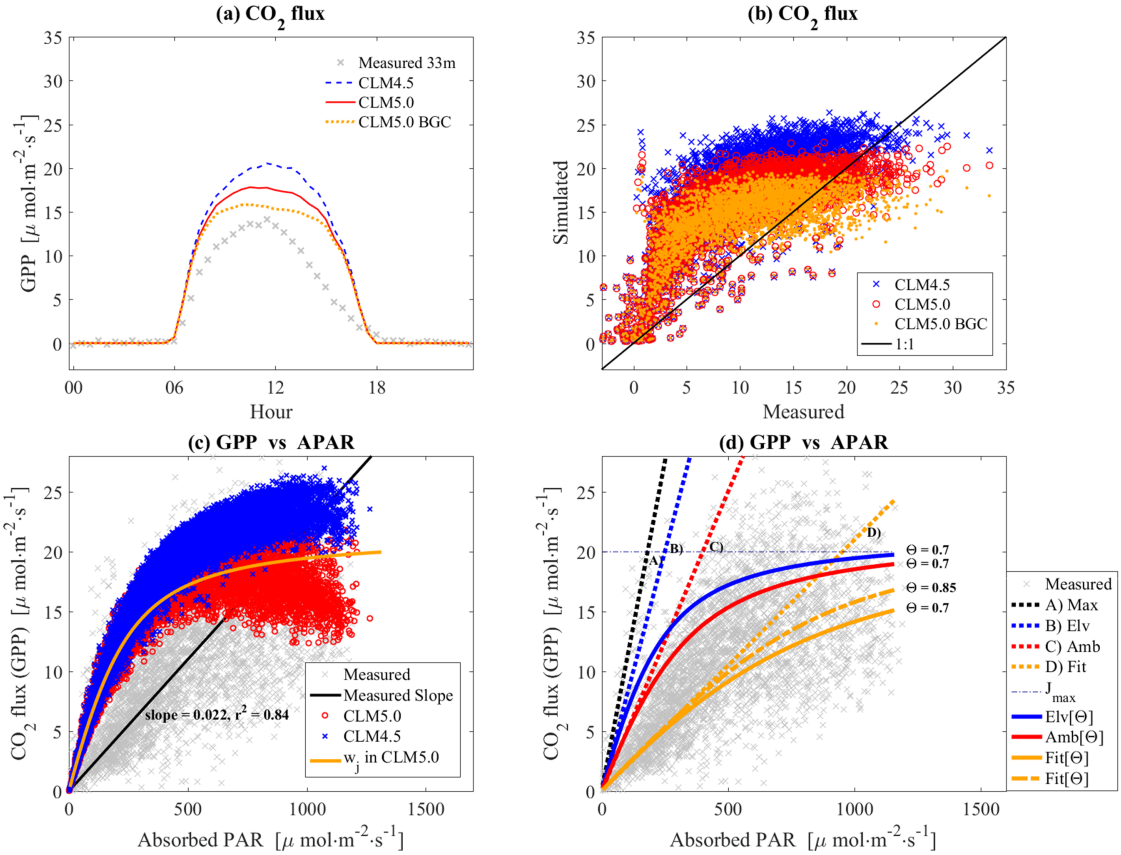

Figure 4.4: (a) The ensemble diurnal variation of CO_2 flux: differences between Eddy Covariance (Canopy Tower 33m) and CLM in daytime is 6.7, 4.9, and 3.6 μmol·m^{-2}s^{-1} for CLM4.5, CLM5, and CLM5BGC, respectively; (b) One to one plot in reference to data shown in figure (a); (c) APAR vs GPP, and w_j is simulated with default parameters and $\mathbb{C}_i$=0.667; (d) 'A)' is the theoretical maximum $\alpha\approx0.11$, 'B)' is saturated/elevated condition $\alpha\approx0.075$ (absence of photorespiration), 'C)' is normal atmosphere condition $\alpha\approx0.05$ if $\Phi\approx0.6$, and 'D)' is a fitted value $\alpha\approx0.021$ if $\Phi\approx0.25$ from our observation. [Θ] means applying the hyperbolic function Eq. (4.7) with the curvature value Θ. The change of the slope can alter maximum assimilation rate, and the alteration can be counterbalanced if Θ is modified.

4.3.3 H_2O Flux

The effect of the change of f_{wetmax} can be detected in the model's results for vapor fluxes [Figure 4.6a; Figure 4.6c; Figure 4.6e; Figure 4.6f]. In this simulation, since CLM5 has a low leaf

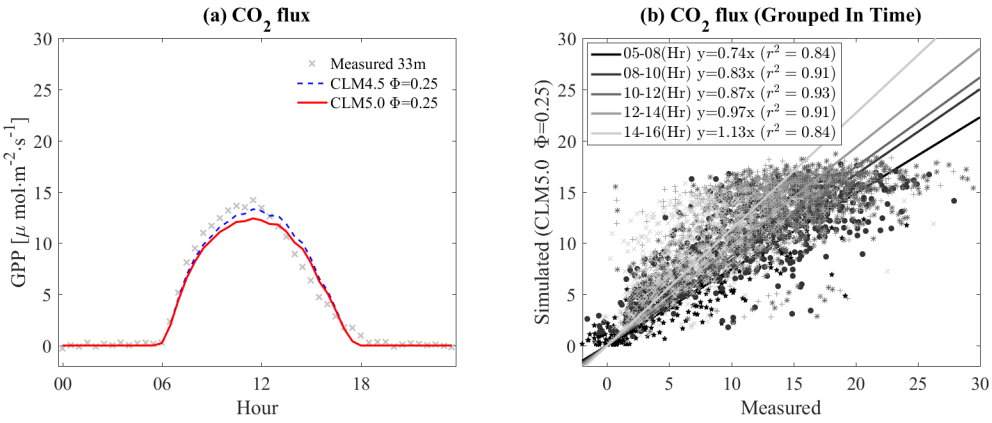

Figure 4.5: Test simulation using Φ=0.25 and Θ=0.7 as in [Figure 4.4d]. By the modification of Φ, maximum GPP has been reduced. It is possible to improve this model by updating Θ. r^2 is a R-squared value without an intercept.

wetness coefficient (i.e., maximum rate is 0.05 as Eq. (4.4), which reduces canopy evaporation and elevate transpiration rate), f_{wetmax} was considered as 1 for CLM5 and it is hereafter named as (CLM5 fmx=1).

Similarly to CO_2 flux, total H_2O fluxes of CLM5 are overestimated (2.1×10^{-5} $mm\cdot s^{-1}$ higher in daytime than Eddy-Covariance). Flux rate in CLM5 with fmx=1 which is for fair comparison is reduced in comparison to the previous version [Figure 4.6a; Figure 4.6b]. The notable decrease (CLM4.5 & CLM5 with Φ=0.25) was due to the change of the quantum yield α parameter as done for the GPP simulation [Figure 4.5], which results in the further reduction but in the case of CLM5 still 1.1×10^{-5} $mm\cdot s^{-1}$ higher than measurement during daytime. Meanwhile, transpiration rates (TR) are overestimated in both CLM4.5 and CLM5 (Figure 4.6d) but there is also notable improvement when changing the quantum efficiency parameter, as done previously for GPP and total vapor flux. At the daily time scale, CLM4.5 produces the highest estimates for both ET and TR in comparison to the other versions. CLM5 yields a notable and reduced change to the estimated ET and TR due to the newly implemented leaf wetness parameter f_{wetmax}. Applying a quantum efficiency of Φ=0.25 makes fitted lines more close to the 1:1 line in both ET and TR [Figure 4.6e; Figure 4.6f]. Here, correlations of TR are slightly increased by around 0.01 ($R^2_{CLM4.5}$ = 0.67, R^2_{CLM5} = 0.68) when considering Φ = 0.25, however, the correlations of ET are decreased

by around 0.1 ($R^2_{CLM4.5}$ = 0.42, R^2_{CLM5} = 0.44) [Table 4.1]. The reduced TR when assuming a lower quantum efficiency makes the fitted slope for ET decrease [Figure 4.6e], possibly since transpiration rate is a more influential component than evaporation in this site. Thus, driving ET rates when there is higher energy exchange condition (i.e., warm, sunnier and drier time). The daily-scale comparison which is up-scaled except nocturnal transpiration as mentioned previously, If one wants to include nighttime transpiration, the all intercept values (y-axis) in [Table 1] for transpiration are decreased by -2.2$\times 10^{-6} \cdot mm \cdot s^{-1}$.

Table 4.1: Fitting parameters and regression coefficients for sap flow and eddy-covariance measurements (except nighttime transpiration) versus simulation by CLM in daily scale for [Figure 4.6e] and [Figure 4.6f].

Figure	**Line**	**Model**	**Data**	**Slope**	**Intercept** ($10^{-6} \cdot mm \cdot s^{-1}$)	R^2
[Figure 4.6e]	1	CLM5	EC 33m (ET)	0.92	9.98	0.51
	2	CLM5 fmx=1		0.82	10.79	0.51
	3	CLM4.5		1.04	8.06	0.55
	4	CLM5 Φ=0.25		0.69	8.51	0.42
	5	CLM4.5 Φ=0.25		0.75	7.65	0.44
[Figure 4.6f]	1	CLM5	Sapflow (TR)	1.01	-0.22	0.66
	2	CLM5 fmx=1		1.17	-5.32	0.67
	3	CLM4.5		1.37	-7.29	0.67
	4	CLM5 Φ=0.25		0.79	-2.14	0.67
	5	CLM4.5 Φ=0.25		1.03	-5.84	0.68

4.3.4 Leaf Wetness

Simulated leaf wetness via CLM is not in agreement with the diurnal leaf wetness variation measured in this site [Figure 4.7b]. In particular, the night-time fraction of leaf wetness is significantly higher when compared with gauged data. The biggest problem detected in this study is that intercepted canopy water is rarely evaporated in the model, especially due to frequent nighttime rainfall (starting in late afternoon) or high daytime humidity (condensation). Daytime leaf wetness seems to be reasonably simulated [Figure 4.7b], but the comparison between simulated and mea-

sured data (not displayed here) cannot identify any trend, which indicates that the formula cannot adequately represent actual behaviors of wet fraction in both CLM5 and CLM4.5.

Secondary analyses show that leaf wetness behavior is highly sensitive to incoming solar radiation. The analyses are to track the behavior of drying leaf in [Figure 4.7e; Figure 4.7f], which is drawn via the collection of consecutive no-rain events for 3 hours just right after last rain events in the daytime between 10 AM and 14 PM. Although it is difficult to gather such a serial drying event (each plot uses at least 12 successive datasets), the result clearly shows that leaf wetness is strongly influenced by increase of incoming solar radiation when f_{wetmax}=1 (CLM5 fmx=1 and CLM4.5). In the case of f_{wetmax}=0.05 (CLM5), the drying rate is reasonable at low solar radiation, but it is still higher than observed with high incoming solar radiation. The measured data in the analysis shows relatively small values of leaf wetness at lower levels of the canopy, which indicates that the lower canopy does not sensitively respond to the interception for rain. This aspect can be a case to support lower f_{wetmax} can be reasonable.

Intercepted precipitation is usually too high in CLM compared with observed leaf wetness [Figure 4.7c; Figure 4.7d]. The analysis made via detecting the increment of the leaf wetness by precipitation, which is directly related with canopy interception: the usual increment for 2-hour (30-min) rain is 0.71(0.33) at 38m, 0.48(0.28) at 3m observation, around 0.88(0.73) in CLM5, 0.97(0.77) in CLM5 fmx=1, and 0.94 (0.46) in CLM4.5. The interception model in CLM5 fmx=1 has been updated in Eq. (4.3) resulting in higher interception rate than CLM4.5 fmx=1 [Eq. (4.2)]. Also, the interception rate seems higher with CLM5 fmx=1 than with CLM5 as in [Figure 4.7c]. This is because CLM5 fmx=1 has higher canopy evaporation rate which results in the acceleration of canopy evaporation while allowing interception to assume a larger role in the canopy water balance. In the one to one comparison, the increase of leaf wetness in CLM is usually higher than in measured data. Consequently, the wet canopy fraction at the beginning of the drying process right after the last rain events is usually higher in CLM than in measurement: 0.63 at 38m, 0.47 at 3m observation, 0.96 in CLM5, 0.9 in CLM5 fmx=1, and 0.78 in CLM4.5 (see y-axis data at 0 x-axis in [Figure 7e]).

4.3.5 Temperatures and Soil Flux

The simulated canopy air temperature in both CLM4.5 and CLM5 is overestimated during daytime (+0.8 and +1 °C, respectively) and underestimated during the nighttime (-1.9 and -1.1 °C, respectively) as seen in the ensemble diurnal variation analysis [Figure 8]. In other words, the model results have more variation than measured data, and the simulated temperature might be oversensitive to incoming solar radiation, as seen for leaf wetness. Updated MOST parameters improved nighttime air temperatures in CLM5 (Burns et al., 2018), but they are still underestimated. Remaining water on the canopy during nighttime tends to not be efficiently evaporated as reported in the previous section [Figure 4.7b] which is also possibly related to low canopy temperature in CLM. At lower canopy levels, the ground air temperature at the surface is also overestimated during daytime, and it is even higher than air temperature at 1m-5m [Figure 4.8b].

Ground surface tends to have high energy exchange during daytime like as canopy process, compared measured data. Considering the soil temperatures [Figure 4.8] and the soil heat fluxes [Figure 4.9], we found they are overestimated during daytime and underestimated in the nighttime. Soil temperature and heat flux in CLM was highly variable, as expected due to results from previous sections. Soil evaporation rates in both CLM4.5 and CLM5 are also overestimated compared with estimated data from soil respiration chamber measurements (LI8100) [Figure 4.9]. For soil evaporation, averaged differences in daytime is 5×10^{-7} $mm{\cdot}s^{-1}$ with CLM4.5 and 15×10^{-7} $mm{\cdot}s^{-1}$ with CLM5, while measured field value is around 1×10^{-7} $mm{\cdot}s^{-1}$. The simulated soil moistures also have high changeability with low mean water contents (around 0.2 $m^3{\cdot}m^{-3}$) compared with gauged values (0.3-0.4 $m^3{\cdot}m^{-3}$).

The overestimation of vegetation temperature (Tv) in both CLM4.5 and CLM5 also appears in the daytime ($\approx$ + 1.0$\sim$2.4 °C) through the ensemble diurnal variation and the one to one comparison [Figure 4.10a; Figure 4.10b]. Another model test is also made using global forcing datasets (Qian et al., 2006) to corroborate our simulations, and the result was a very similar behavior ($\approx$ + 5 °C, not depicted here). The high Tv and Ta from CLM simulations results in lower relative humidity than gauged-based canopy air humidity (figures are not presented here). We need to note that

the sunlit/shade scheme in CLM does not consider two different vegetation temperatures, so it only takes single variable Tv representing entire canopy. Canopy temperature (Tv) in CLM should be an average of sunlit and shaded leaf temperature but the simulated results are far from our expectation [Figure 4.10a]. A comparison plot also shows significant error [Figure 4.10b]. The additional comparisons indicate that Tv on sunlit leaves normally follows the canopy air temperature (leaf thermoregulation) but CLM does not have such properties [Figure 4.10c; Figure 4.10d].

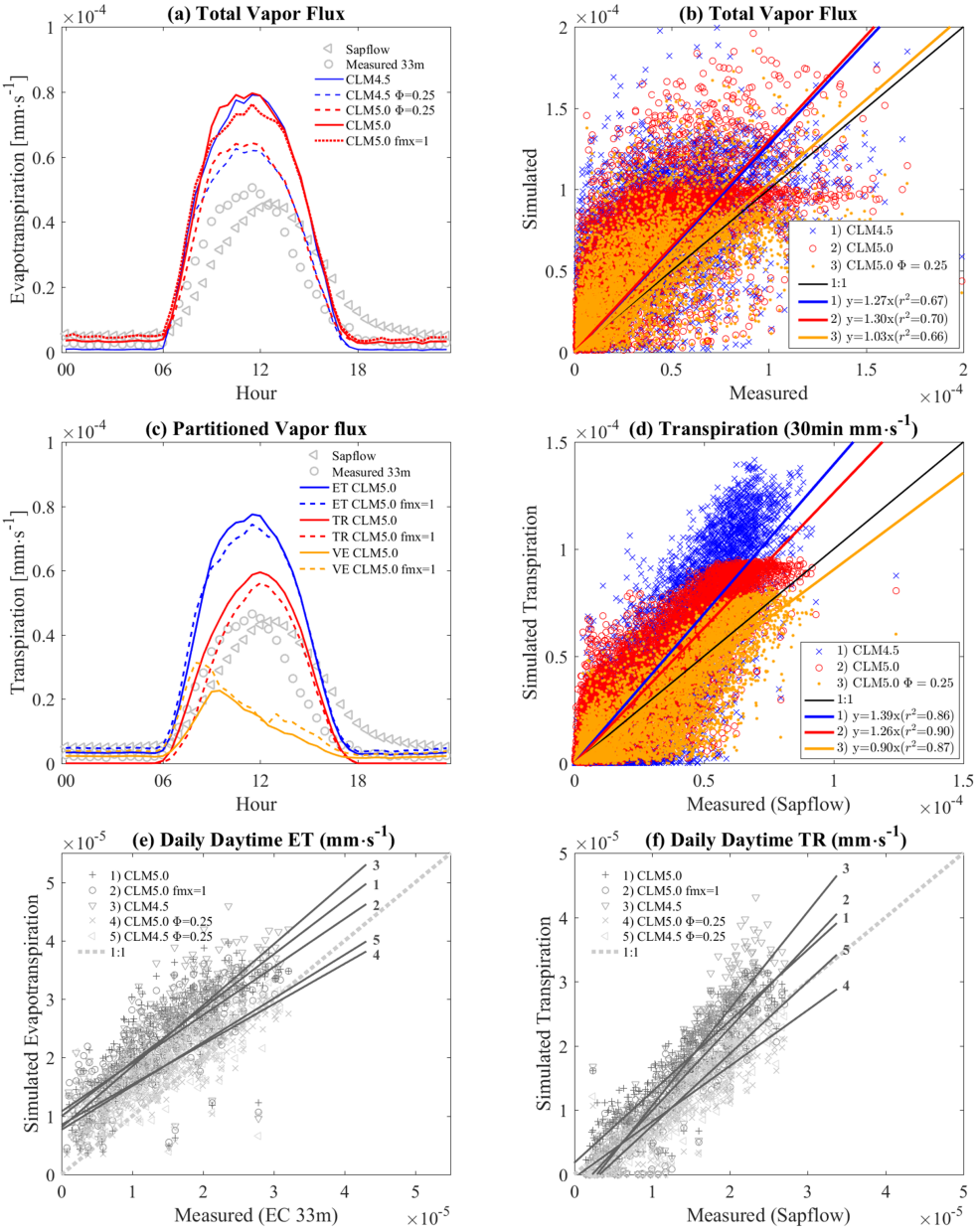

Figure 4.6: (a) The ensemble diurnal variation of total H_2O flux, where "Measured 33m" is measured by Eddy Covariance (at 33m), "Sapflow" is transpiration measured through sapflow, all "CLM" are about evapotranspiration (ET), 'fmx=1' sign represents fwetmax=1, and 'Φ=0.25' means 0.25 applied to Φ in Eq. (4.7); (b) One to one plot in reference to data shown on figure (a); (c) diurnal variation of partitioned H_2O flux, where ET, TR, and VE are evapotranspiration, transpiration, and canopy evaporation from CLM; (d) The one to one plots of 30 min scale transpiration rate between simulated TR versus sapflow transpiration rates; (e) The one to one plots of daily evapotranspiration (except nighttime transpiration) rate between simulated ET versus measured one from eddy-covariance; (f) The one to one plots of daily transpiration rate between simulated TR versus sapflow data (except nighttime transpiration)

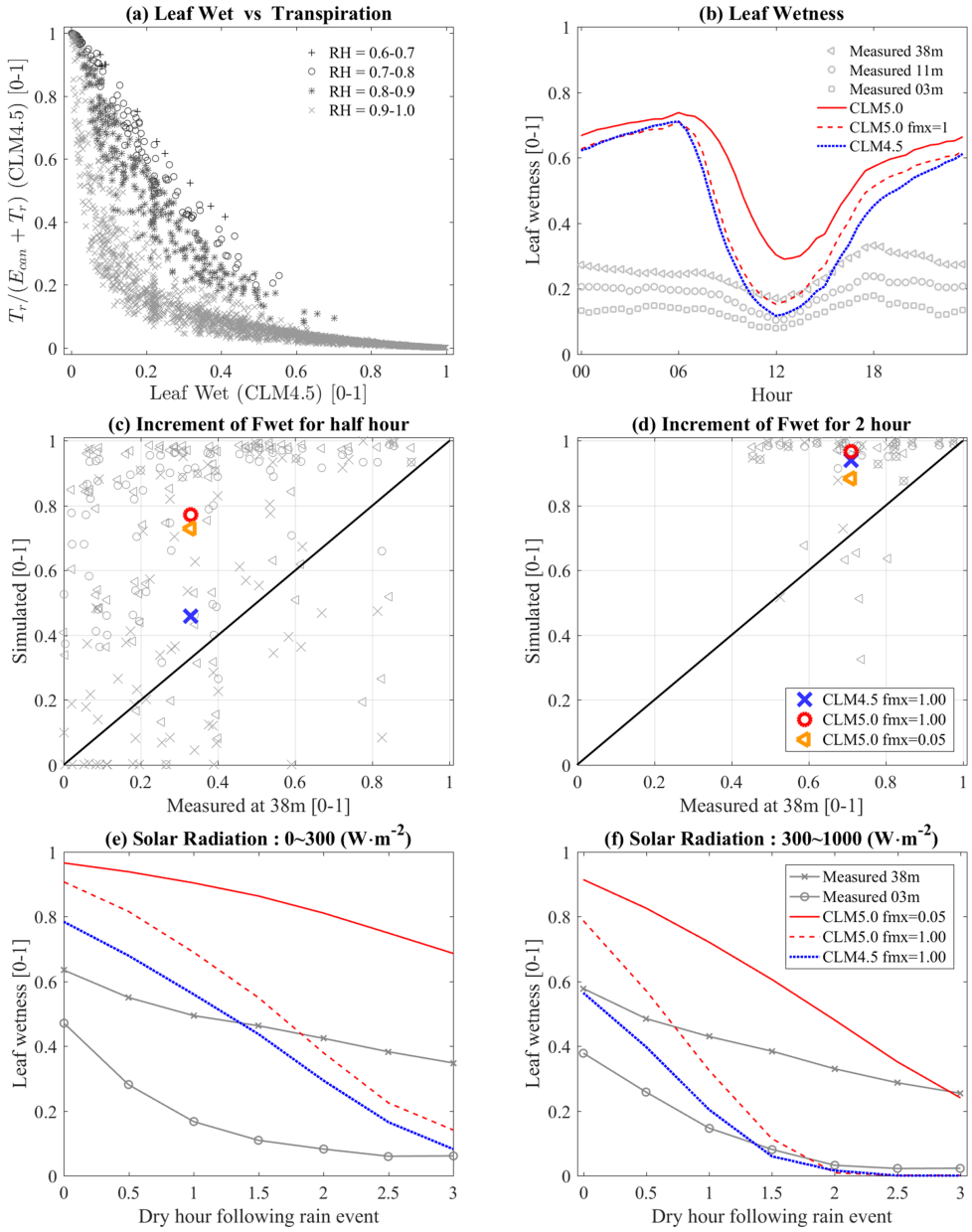

Figure 4.7: (a) Transpiration/Evapotranspiration versus leaf wetness and classified by relative humidity [0-1]; (b) The ensemble diurnal variation of leaf wetness, where '38m', '11m', and '3m' are measurement heights and the others are leaf wetness from CLM5 (f_{wetmax}=0.05), CLM5 fmx=1 (f_{wetmax}=1), and CLM4.5 (f_{wetmax}=1); (c) and (d) are the behavior of wetting canopy at 38m and large markers are the average of each simulated value. The collected data has the condition that there is no previous rainfall event at least in 2 hours and leaf wetness is lower than 0.2 at the beginning; (c) is for 0.5-hour rainfall events (one consecutive event in 30-min scale); (d) is for 2 hours rainfall events (four consecutive events); (e) and (f) are the behavior of drying canopy. For this analysis, Consecutive no-rainfall (3 hours) data and daytime between 10 am and 2 pm are collected right after rain event. Marked lines are from measurements, and continuous lines are estimated from CLM; (e) is when solar radiation is low (0 - 300 $\mathrm{W{\cdot}m^{-2}}$); (f) when solar radiation is higher than 300 $\mathrm{W{\cdot}m^{-2}}$. All leaf wetness [0-1] from CLM was re-estimated based on Eq. (4) because such CLM5 has a limit as f_{wetmax}=0.05. Also, whether or not to apply the power of 2/3 does not change much our comparisons.

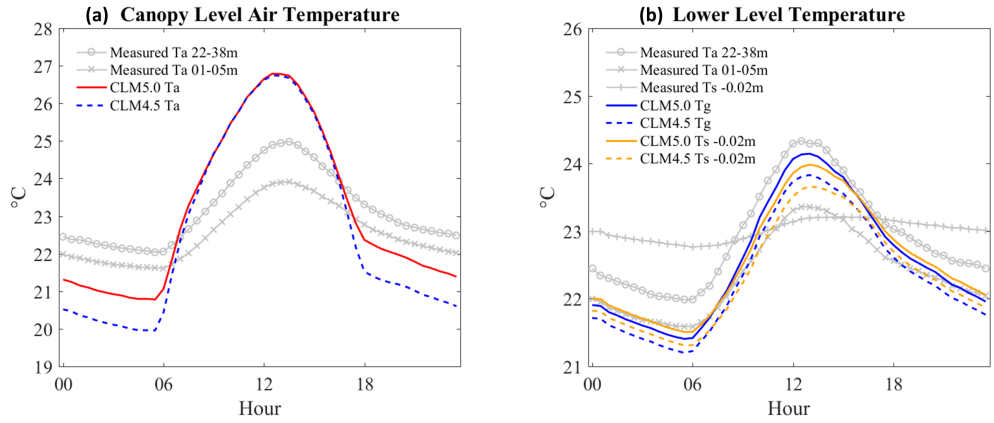

Figure 4.8: The ensemble diurnal variation of air temperatures. Canopy Levels at 22-33m and 1-5m are measured air temperatures, Ta represents air temperature at 28.075m in CLM, and Tg is ground air temperature at 0.01m in CLM. 'Ts -0.02m' is measured/simulated soil temeprature. In (a), both CLM4.5 and CLM5 is overestimated in daytime (+0.8 and +1 °C) and underestimated during the nighttime (-1.9 and -1.1 °C). In (b), differences bettwen 'Measured Ta 01-05m' and all CLM values (CLM5.0 Tg, CLM4.5 Tg, CLM5.0 Ts, and CLM4.5 Ts) are -0.39, -0.14, -0.32, and -0.06 in daytime and -0.02, 0.18, -0.11, and 0.08 °C in nighttime. Differences with 'Measured Ts -0.02m' are -0.04, 0.21, 0.03, and 0.30 in daytime and 0.90, 1.10, 0.81, and 1 °C in nighttime.

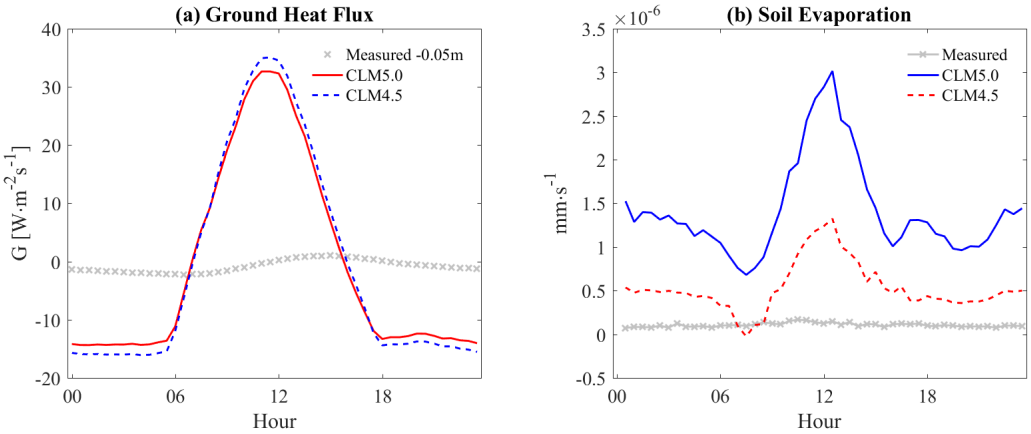

Figure 4.9: The ensemble diurnal variation of soil/ground heat fluxes (into soil +) (left) and soil evaporation.

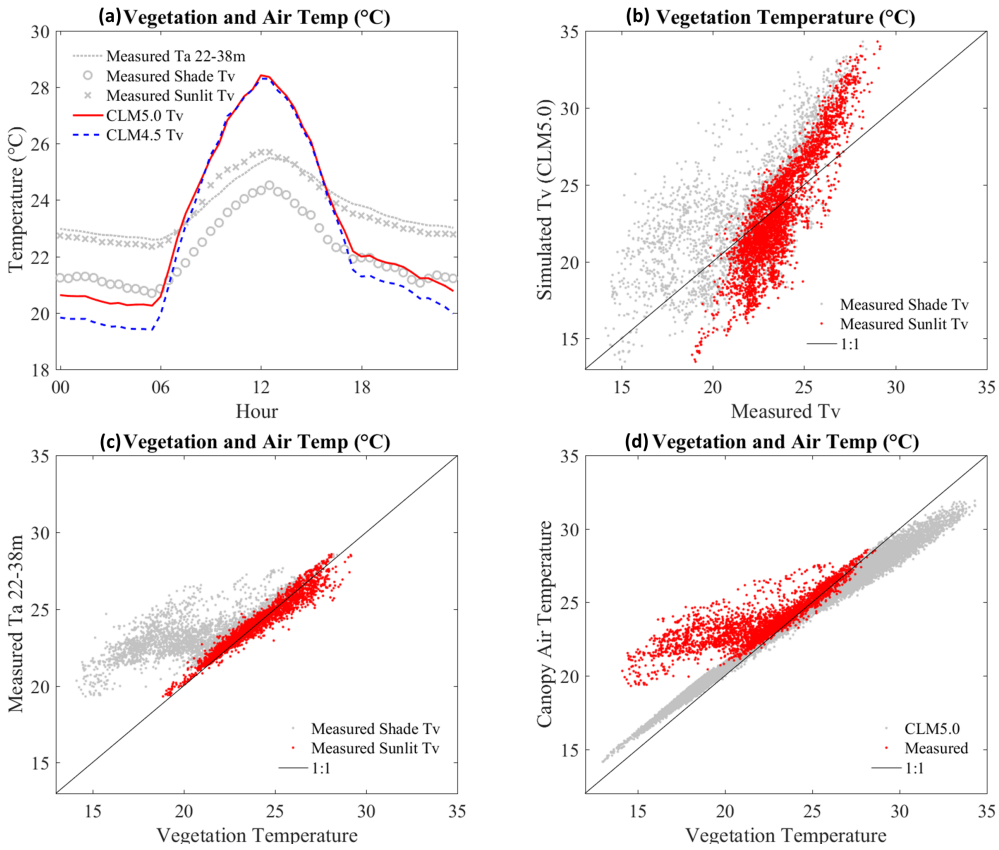

Figure 4.10: (a) The diurnal variation of leaf temperatures with measured canopy air temperatures; (b) The one to one plot of leaf temperatures: CLM vegetation temperatures (Tv) are compared with measured values for the both gauged shade (Shade Tv) and sunlit (Sunlit Tv) vegetation temperatures; (c) The one to one plots about measured canopy air temperatures versus measured leaf temperatures (sunlit and shade); (d) The one to one plots about canopy air temperatures versus leaf temperatures from CLM (CLM5 Ta vs CLM5 Tv) and observation (Canopy Ta 22-38m vs averaged Tv from sunlit and shade Tv): Averaged Tv is estimated through $(LAI_{Shade} \times Tv_{Shade} + LAI_{Sunlit} \times Tv_{Sunlit})/LAI$. In (a), daytime differences 'CLM 5.0 Tv' minus measurments ('Measured Ta 22-38m', 'Measured Shade Tv', and 'Measured Sunlit Tv') are 1.1, 2.4, and 1.0. In nighttime, the differences are -2, -0.3, and -1.8 °C. CLM5 normally 0.2 higher in daytime and 0.8 higher in nighttime.

4.4 Discussion and Conclusion

In this study, two versions of the Community Land Model (CLM4.5 and CLM5), running primarily in the satellite phenology (SP) mode, were tested against measured data from a mountainous tropical rainforest in Costa Rica. Net radiation was underpredicted by both CLM4.5 and CLM5, by an average of -20 $W \cdot m^{-2}$ [Figure 4.3a; Figure 4.3b]. The discrepancy was attributable to CLM's over-prediction of surface albedo, which showed a +0.022 difference [Figure 4.3c; Figure 4.3d].

The effects of topological slope clearly appear in the diurnal plots for albedo/PAR [Figure 4.3] and for CO_2 flux [Figure 4.5]. With respect to albedo, the hillslope shading effects magnified these discrepancies, with afternoon values having larger differences as the sun moved behind the north-south trending mountain [Figure 4.3]. Different degrees of discrepancy exists according to the diurnal cycle of the intensity of incoming solar radiation and solar angle [Figure 4.3c; Figure 4.3d]. PAR profiles also show radiation levels within the canopy have a skewed cycle [Figure 4.3e; Figure 4.3f], which were not captured by CLM. These indicate that canopy radiative transfer, including the surface albedo and sunlit/shade separation in CLM, may need to represent such characteristic for advanced land surface model to reflect a more realistic response to solar radiation or topographical slope. More importantly, aerodynamic resistance models, such as MOST, are currently incapable of representing a sloped terrain. If such effects can be implemented in CLM, predictions can be highly improved, especially at a fine grid scale.

This study demonstrates the possibility of reducing predictive uncertainty by adapting the model to mimic such slope effects, and suggests that additional observations are necessary to examine and capture such features. Many past studies to compare and improve CLM have taken a similar approach to this study, but with more focus on specific sub-model performance (Burns et al., 2018; Swenson and Lawrence, 2014; Bonan et al., 2011), rather than studying spatial complexity. The slope effects for albedo are minor in this study, because the hysteresis effect (which appears as skewness in the diurnal average curve) is relatively small and because the effects cannot be visually identified from the net radiation curves. On the other hand, the skewness for PAR is significant and this is obviously related to the different response of GPP through time [Figure 4.5].

Such influence might not be noticeable if the GPP comparison were not classified by time, because the error appears similar to white noise. If this effect is captured, the prediction of physiological variables (e.g., GPP and TR) can be improved. We anticipate the same effect would be present in a wider range of forests. Also, recent land surface models are becoming more elaborate vertically (e.g., multi-layered model (Bonan et al., 2018; Ryder et al., 2016)) and horizontally (e.g., vegetation demographics (Fisher et al., 2018), and the performance of these advanced models would be affected by topographical characteristics. Hence, further investigation should focus on both improved model parameterization for hillslopes and additional data from mountainous forests.

The simulated photosynthesis rate tends to be higher than those observed; these result are also reported in similar studies of montane rainforest (Fan et al., 2019; Muñoz-Villers et al., 2012); such errors could possibly be alleviated by updating parameters associated with light-limitation effects. For carbon flux (GPP) and transpiration (TR), the over-estimation in CLM4.5 has been reduced in CLM5 [Figure 4.4b; Figure 4.6b; Figure 4.6d]. However, CLM5 and CLM-BGC seem to reduce the maximum assimilation rate limit by lowering the BB slope photosynthesis parameters (i.e., $V_{c,max25}$ or J_{max25}). The curved-shape error in GPP, the middle range of photosynthesis rate, still exists compared with CLM4.5 [Figure 4.4b], and one of the light-response photosynthesis model could be the cause. We have briefly addressed the electron transport model (Eq. (4.6) and Eq. (4.7)) and tested it by changing quantum yield and curvature parameters [Figure 4.4c; Figure 4.4c]. The GPP including transpiration shows the change of the parameter (fitted quantum yield) results in better agreement with observation and additionally a clear effect of the topographical slope appears [Figure 4.5; Figure 4.6f].

Partitioning the water flux is a critical task which also needs more investigation. Errors in vapor flux are particularly difficult to diagnose since the discrepancy can be caused by the failure of any of the embedded sub-models, although transpiration is the largest driver of the overall pattern of total vapor flux (ET) [Figure 4.6c]. For comparison, CO_2 flux in CLM largely depends on plant-light relationships (photosynthesis). In contrast, evapotranspiration (ET) consists of three major components: soil evaporation, canopy evaporation, and transpiration. Therefore, an error in any

one of the sub-models can make the entire water flux (ET) inaccurate. We can also recognize that the comparison of total vapor flux [Figure 4.6b] has much uncertainty compared to CO_2 flux [Figure 4.4b], which has a strong curved correlation at least

Canopy evaporation, which relies on both the rainfall interception sub-model and the leaf wetness sub-model, is a key to proper partitioning. Both ET and TR are affected by the Canopy evaporation [Figure 4.6a; Figure 4.6c], because leaf wetness suppresses transpiration and enhances canopy evaporation in CLM [Figure 4.2b; Figure 4.7a]. However, the leaf wetness variable in CLM has high uncertainty in a number of analyses, including ensemble diurnal variation [Figure 4.7b] and interception rate [Figure 4.7c; Figure 4.7d], possibly due to oversimplifaction of throughfall processes as reported in a previous study (Fan et al., 2019). Leaf wetness related parameters are optimized for large-scale forcing (e.g., 6 hourly data). The improperly modeled canopy water levels and the wetted fraction result in errors in canopy evaporation which overreacts to the intensity of solar radiation (or net radiation) [Figure 4.7e; Figure 4.7f]. We observe some improvement in CLM5 by low maximum wetness f_{wetmax} but the simulated leaf wetness is still sensitive to the incoming solar energy. Such water-related process can also have vertical/spatial variation due to the structure and the shape of canopy but also the topographical effects.

The new maximum leaf wetness applied in CLM5 may need to vary more by vegetation type, as highlighted in a previous study (Fan et al., 2019). Changing f_{wetmax} has a significant impact on latent heat fluxes [Figure 4.6a; Figure 4.6c; Figure 4.6e; Figure 4.6f], contrary to the results noted by Burns et al. (2018). While this affect could be attributed to much more frequent rainfall at our site, it also could be that a low f_{wetmax} is more reasonable for needle leaf species than it is for those with large, broad leaves. Leaf surfaces in canopy f_{wet} cannot be easily fully-wetted even with conifer, and transpiration rate is not much sensitive with relatively high leaf wetness ($f_{wet} > 0.4$) [Figure 4.7a], however, simply applying f_{wetmax} as 0.05 for all sites cannot be realistic. The role of leaf wet faction is not negligible in CLM, and the behavior of the photosynthesis is still sensitive to leaf wetness ($f_{wet} \leqq 0.4$). At low relative humidity values, the relationship between leaf wetness and transpiration is close to linear, but it tends towards non-linearity at high

relative humidity [Figure 4.7a]. In our site, different leaf wetness behaviors have been observed between sunlit canopy and shaded canopy (Andrews et al., 2019), which can be also an important characteristic for tracking canopy evaporation and ET.

Temperature-related variables are also problematic in CLM [Figure 4.8; Figure 4.9; Figure 4.10], which may be caused by errors in energy partitioning and modeling aerodynamic resistance. Day vs. night changes in canopy air temperature and leaf temperature in CLM are excessive, an issue also noted in previous comparisons. In the Burns et al. (2018) study, changing the MOST parameters partially corrected underestimates of nighttime air temperature in CLM5. The source of error is similar to, and perhaps intertwined with, the issues found with leaf wetness. Consequently, because the canopy air temperature is linked with soil/ground temperature, following soil temperature and soil heat flux including the ground evaporation in CLM also has higher fluctuation than measured datasets [Figure 4.9]. These variables (soil moisture, soil temperature, and soil heat flux) are highly related each other and they are also linked with the canopy condition, so it is difficult to conclude which part most results in such high variation in the soil system. Additionally, the CLM results have low relative humidity at canopy air (lower than measured data) in daytime. This is also possibly related to the overestimated canopy air/leaf temperatures, which can also affect physiological simulation such as transpiration rates. Such overresponse to incoming solar radiation can be caused by the failure of energy partitioning and the estimation of aerodynamic resistance, but it is challenging to attribute the error precisely. For instance, ET is overestimated in the daytime, which results in low thermal exchange between the surface and the atmosphere according to the energy balance. On the other hand, the temperature gradients between the atmosphere, the canopy, and the ground are higher than anticipated [Figure 4.8], which is possibly triggered by the error of aerodynamic parameters. We need to note that the aerodynamic resistance is likewise affected by the air temperature gradient so they are mutually dependent.

Adjustments in light-related parameters (e.g., LAI, leaf angle, and optical depth) did not noticeably improve model results. The ratio of the absorbed energy on the soil surface to the total incoming solar radiation in CLM is 0.0342, but our PAR profile data [Figure 4.3e] indicates the

ratio should be lower, around 0.01. Even though the modeled ground surface tends to receive excess solar energy, changing this value does not seem to result in significant improvement in any simulated variables, because it is a relatively low portion of the energy budget. Likewise, increasing LAI to 7.7, based on nearby site measurements (Teale et al., 2014), only slightly alleviated issues associated soil temperatures and made no difference in canopy temperatures. We have also tested with different leaf angles which are directly related to the optical depth (K), but there is no significant difference; a change in leaf angle from χ_L=0.1 to χ_L=0.4 results in a 0.3 °C decrease in ground surface temperature. These supplementary tests indicate that the reduction of absorbed solar radiation on the ground and the some changes of parameters for soil albedo do not significantly alter canopy temperatures. The problem may more likely be caused by errors in the aerodynamic resistance above the canopy or oversimplified canopy structures, as has been reported in other studies (Wang et al., 2014; Chen et al., 2010; Zheng et al., 2012; Zeng et al., 2012).

A complete multi-layer scheme may be necessary to improve the model. From the Tv analysis, we find that both CLM5 and CLM4.5 have two big-leaf scheme as sunlit/shade area in the canopy. However, this module works only for incoming solar radiation, not for leaf temperature or canopy air temperature. The air temperature differences along heights within the canopy are not significant based on measured data, but leaf temperature can be a promising variable to be partitioned into sunlit and shade temperatres due to that the two have somewhat different behaviors: see between measured sunlit and shaded vegetation temperatures (the fraction of sunlit is about 26% in CLM) [Figure 4.10]. The fraction of leaf wetness also represents the entire canopy area in CLM, which seems physically too simplified. The higher location of canopy tends to be easily wetted/dried than the lower location; the more exposed canopy area (higher location) is normally wetter than shaded canopy area [Figure 4.7e; Figure 4.7f].

Vegetation temperature affects energy flux temperature via its relationship to canopy air temperature (Ta) and physiological processes such as transpiration (Wang et al., 2014). Nevertheless, the problem of skin (surface/leaf) temperature appears in this study as other reports (Wang et al., 2014; Chen et al., 2010; Zheng et al., 2012; Zeng et al., 2012). Some researchers attribute these

issues to incorrect parameterization of roughness length for heat and have made a number of advances toward reducing these errors (Yang et al., 2002; Wang et al., 2014; Chen et al., 2010; Zheng et al., 2012; Zeng et al., 2012). However, we note that our case is different since most studies discussed low diurnal variations and underestimations. The one to one comparisons between the canopy air temperature and the leaf surface temperature [Figure 4.10c; Figure 4.10d] indicate that Tv on sunlit leaves normally follows the canopy air temperature (leaf thermoregulation) as described in other literature (Michaletz et al., 2016). However, CLM does not consider such leaf thermoregulation processes.

In conclusion, through the simulation test of CLM on land-atmosphere processes of a steep mountainous tropical rainforest, we identify how the global-scale parameterization works at this unique site. These are very few case studies of such steep surface and wet tropical forest, which more highlights the importance of this study. We have found that CLM5 has some advantages over CLM4.5. However, CLM5 does not yet sufficiently resolve the critical problems (such as partitioning the energy balance), and updates to the representation of in-canopy processes (such as, photosynthesis model, turbulence model, and model structures) are still needed to capture temperature variations and physiological activity. More importantly, further investigation into including slope effects into the models is required.

Some detailed results are that canopy temperatures and leaf temperatures are over-sensitive to incoming solar radiation, which can cause a number of issues: low relative humidity near canopy surface, subsequently affecting tree physiological processes, and excessive heating of the soil surface, leading to unrealistically high average and day-to-night differences in soil temperatures and soil heat fluxes. The formulation describing leaf wetness processes is too simplified, causing model failure for the frequently rainy areas. Transpiration rate, which is the largest part of latent heat flux at the site, as well as carbon uptake (photosynthesis activity) are also over-estimated in CLM. We have also suggested potential improvement of photosynthesis model by the re-parameterization of quantum efficiency. Ultimately, however, it may be necessary to apply a two big leaf scheme or multi-layer scheme to better depict the multi-faceted interactions between leaf wetness, temper-

ature, and shading to properly represent canopy processes in tall, dense, or mountainous forests such as this.

Based on these new findings, the application and further investigations are necessary. In particular, actual improvement at this study site by applying new parameterizations and global-scale tests will be the next goal. Also, more observations at both this site and other new places are essential to increase the number of samples and to enhance the reliability of the land surface model.

5. PROFILE MODELING OF MICROMETEOROLOGICAL VARIABLES IN A TROPICAL MOUNTAIN RAINFOREST USING MULTILAYERED CLM (CLM-ML)

5.1 Introduction

Land surface models (LSM) do not yet adequately capture land-atmosphere interactions (Cai et al., 2014; Wang et al., 2014; Lawrence et al., 2011; Oleson et al., 2010). In particular, as discussed in the previous section, a LSM prediction is likely contains significant errors in tropical regions (Bonan et al., 2011, 2012), although the LSM such Community Land Model (CLM) has been updated for continued improvement.

One of the main reason of model failure discussed in previous section is that a single layer model can be insufficient to represent various surface environments. Consequently, inter-comparison study or site-based study based on the single layer model (e.g., CLM4.5, CLM5) cannot be adequate for a notable improvement of model. The surface layer is normally described as one or two uniform blocks including soil and vegetation, which is also known as the Big-Leaf model (Dai et al., 2004; Oleson et al., 2013). Applying such as model to the energy budget in the surface layer provides an efficient and tolerable degree of accuracy for the many study sites (Ryder et al., 2016). However, it cannot sufficiently represent the response of various fluxes from the surface caused by large differences in the generating mechanisms such as vegetation growth, turbulent transfer, and energy exchange within the canopy. The main fluxes such as latent and sensible heat flux 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 and Gao, 2010; Ryder et al., 2016). Some studies also report that Big-Leaf scheme cannot 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 is the failure of the partitioning of incoming solar radiation as function of height which is one of the most vital

inputs to simulate transpiration (TR), carbon uptake, and energy balance (net radiation) through the leaf density.

Due to the potential of multi-layered scheme, development is in this area ongoing on a number of multi-layer canopy models in addition to CLM. However, most of them have similar functional forms, so focusing on CLM can be sufficient to represent effects on 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 sub-models, which themselves can have varying methods and levels of complexity. For instance, in the light profile scheme,the Multilayer Canopy-Root-Soil model (MLCan) (Drewry et al., 2010) and CLM-ml (CLMML) (Bonan et al., 2018) use the radiative transfer model described by Campbell and Norman (2012) which has simple mechanical and iterative procedures. CLM-ml (CLMML) can additionally select to use the multi-layered Two-stream approximation method introduced by Bonan et al. (2011), which is mathematical based model (closed-form) and also used in single-layered CLMs.The Atmosphere-Plant Exchange Simulator (APES) (Launiainen et al., 2015) uses a mathematical scattering model from (Zhao and Qualls, 2005). Multi-layer Simulator of the Interactions between a Coniferous stand and the Atmosphere (MuSiCA) (Ogée et al., 2003) and Organizing Carbon and Hydrology In Dynamic Ecosystems – CANopy (ORCHIDEE-CAN) (Ryder et al., 2016) follow the method in Gu et al. (1999) for a light profile which is different version of the Two-stream radiative transfer model. Simultaneous Heat and Water model (SHAW) (Flerchinger, 2000) uses a radiative transfer model introduced by Flerchinger et al. (2009) which has a similar approach with Zhao and Qualls (2005). For wind profiles or turbulent transfer scheme, MLCan, APES, and ORCHIDEE-CAN employ 1st-order turbulence closure or similar scheme based on K-theory (Massman and Weil, 1999; Katul et al., 2004). On the other hand, MuSiCA, SHAW, and CLM-ml (CLMML) use the extended version of MOST (Harman and Finnigan, 2007; Leuning, 2000). MOST in CLM is widely utilized for many land surface models such as Simultaneous Heat

and Water (SHAW) models (Flerchinger, 2000), Community Land Models (CLM5) (Lawrence et al., 2018), and Community Noah land surface model with Multi-Parameterization options (Noah MP) model (Niu et al., 2011).

When they applied a multi-layered scheme into single-layered CLM, Bonan et al. (2018) also identified the weakness of turbulence model (Monin–Obukhov Similarity Theory, MOST) which has a critical role in determining the amount of each flux such as heat and moisture. The updated turbulence model called Roughness Sub-Layer model (RSL) was successfully applied to the CLM-based multi-layer model (CLM-ml, CLMML) (Bonan et al., 2018). Despite the reasonable predictions of mean gradient and turbulent fluxes, MOST tends to fail within roughness sublayer above or near canopy height (Bonan et al., 2018; Harman and Finnigan, 2008, 2007). By comparing MOST with RSL using CLMML, this study reveals that the turbulent transfer scheme is a key element to determine a model's performance. The study also highlights the change of in-canopy structure in the land surface model produces the reduction of known bias in sensible and latent heat flux, GPP (gross primary production), and also turbulent transfer itself (Bonan et al., 2018).

Despite the update of CLM, a case study with the multi-layer scheme has not sufficiently conducted due to the lack of study sites having profiled observations (Bonan et al., 2018). This is mainly because most sites are initially installed for single-layered scheme/model. Moreover, observations of tropical and geographically complex areas in particular are rare (Song et al., 2019). Case study is important because it can most directly help to diagnose in detail and to develop the more advanced land surface model. Here, the environment of the study site we explore here has several extreme features compared to other sites where data of this type are available (Song et al., 2019), with ambient condition of high humidity, large precipitation, and 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. This study will suggest possible updates to accurately capture the land surface process in such a unique environment.

For this study, in addition to the multi-layered scheme, some different and additional approaches are necessary due to the features of our study site. In a departure from other studies which have normally modeled a flat surface, we will examine the influences of a steep hillslope which mainly affect wind profile (turbulence-transfer profile) and canopy shapes. Tracking turbulence activity is intricate over the sloped surface/canopy because the effect of the hillslope canopy cannot be simply up-scaled to the horizontal boundary of the top grid box, although for flat surface it has been possible that single canopy is regarded as a representative for a study area (grid box). Moreover, to apply in-canopy complexity, two different wind profile methods, (1) RSL scheme embedded in CLM-ml and (2) numerical solution to first-order closure model, are used to estimate the wind forcing value and the wind profile. 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. Leaf area density (LAD) profile and its displacement height (d) are additionally investigated due to site uniqueness. Different LAD distributions are also tested how such difference can affect the model simulation.

In the previous study of this site, some improvements of CLM5 are identified compared to CLM4.5, but the model is not yet sufficient resolved to represent the canopy process of montane tropical rainforest (Song et al., 2019). In this study, another test case is conducted for the multi-layer scheme, and then the model (e.g., wind profile models, canopy shapes, and parameters) is updated to better capture various fluxes and vertical profiles. The detailed information about the study site and those updates are described in a later section.

5.2 Methodology

5.2.1 Micrometeorological Measurements

The study site is the same site in the previous section, located at the Texas A&M University Soltis Center nearby San Isidro de Peñas Blancas in Costa Rica. The site has two primary biometeorological measurement locations. The main weather tower (hereafter called “Met Tower”) is located in a flat clearing over grass at the base of the mountain. The walkup canopy access tower

(hereafter called "Canopy tower") is located within the forest, on the eastern slope. The Met Tower allows for the measurement of meteorological conditions without the influence of canopy processes and resistances.

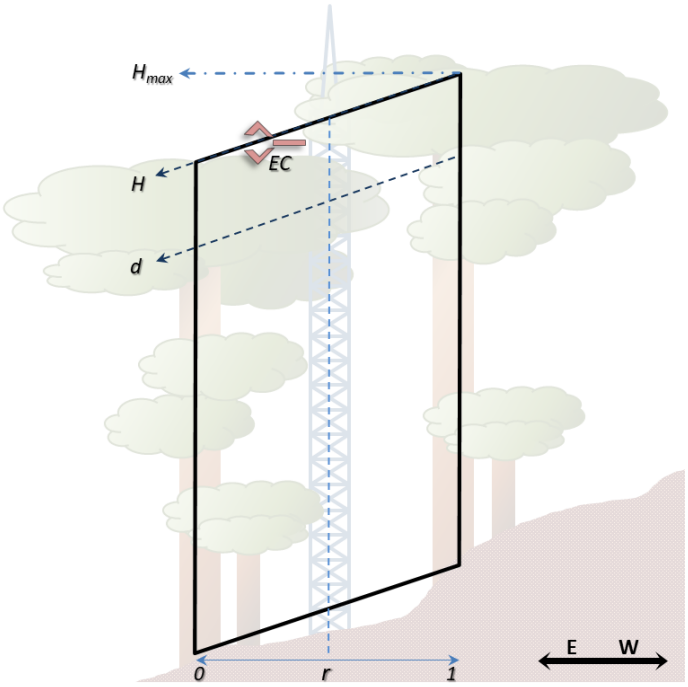

Figure 5.1: Simplified illustration of sloped surface and its two canopies for this study site, where H is average canopy height in large scale, H_{max} is the maximum canopy height at the measurement place, r is the measurement location between two canopies, EC represent eddy-covariance measurement or its sonic anemometer, and d is displacement height.

The Canopy tower is positioned above the canopy but there is an emergent tree [Figure 5.1], leading to a large vertical opening between the canopies from approximately 30 to 40 m. Above the gap, the emergent tree (the upslope tree) provides a substantial degree of sheltering; we see a 70% drop in photosynthetically active radiation (PAR) between downslope canopy surface (30 m) and above the emergent tree (44 m). This configuration can also have some implications which make the eddy covariance method less than ideal. The sonic anemometer and IRGA are located at

33 m, extending away from the tower and clear of obstructions in both the upwind and downslope directions [Figure 5.1]. 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 by the emergent tree (Song et al., 2019).

5.2.2 Forcing Data

As in a previous study which used CLM4.5/5 for comparison with observations, CLMML requires the same forcing data including wind speed, incoming solar radiation, temperature, air pressure, and relative humidity (RH). These data are collected from the Canopy tower in the forest. Missing data are filled with Met tower data, since meteorology data from the two towers are highly correlated with each other. Final missing data are replaced by available data from randomly different day in the same month at the same time, to make the forcing data not to have NaN values. However, these gap-filled data are not compared with observation for this study. Solar data are mostly from the 44 m sensor at the Canopy tower in the mountain forest. Air temperature, RH, and pressure are from the 38 m sensor at the Canopy tower. The air temperature and RH do not much vertical variation (33 m-38m), and these data are also similar to the observation from Met tower.

For wind data, Canopy tower data are used for this study although Met tower is built for the forcing data. The main reason is that it cannot be simply applied as wind forcing due to the effect of the emergent tree. The Met Tower (weather station) collects meteorological data at 10m height over grassland and The Canopy tower also has wind sensors (33 m is the highest). The Canopy tower is placed on the mountain where the actual elevation is much higher (more than 200m by pressure difference) than the Met tower. However, the wind speeds measured in both sites have a similar magnitude: Met tower data (for general forcing data) versus EC data at 33 m tower [Figure B.1]. This means the wind speed at higher location can be much different in the Canopy tower.

5.2.3 Forcing and Canopy Heights

With the previous assumption, the change of forcing height with variables other than wind speed did not give a significant difference for the simulation. In the previous study, we set all forcing heights including the wind as 44 m which is the height of the top radiation sensor. For wind speed, it was mainly because the EC sensor is placed at the lower canopy and above it, as if the radiation sensor is on a high canopy. In this circumstance, due to low displacement height ($\leq$ 35 m), the simulation result was not improved even if we reduced the forcing height to 35 m, which is default height of Canopy in the satellite phenology (SP) mode in CLM. Global forcing data were also used to test this site but yielded similar results, since they have 35 m canopy height and 35m forcing height.

In this study, we assumed that wind data can be affected by two major canopies. In our case, the height of EC for simulation can be influenced by spatially irregular canopy caused by the circumstance of this site. We need to note that the study site is over a sloped surface, and the EC sensor (CSAT3 wind sensor) is located above the lower (downslope) canopy. By the effect of slope, the height of EC at the Canopy tower is 33 m, but its actual height at the lower (downslope) tree location may be higher than that.

Here for various test cases, we use the similar setting of the previous study for all forcing data and also use a higher height (50 m) to start far from the maximum canopy height (43.7 m). At this point, wind speed becomes sensitive to changes in the height. Therefore, it is necessary to predict wind speed at a higher location (50m in this study) based on the measured data which is lower (33 m) than the maximum canopy height (43.7 m by an upslope canopy) among the two trees.

Two test cases for CLMML mimic the previous study about the wind speed forcing, but the other cases in this study mainly use the original height (33 m) of sonic anemometer. CLM4.5/5 simulations and the two test cases of the CLMML set the forcing heights at 44 m as the previous study, to compare each other. Again, this assumes there is no interference from the emergent tree which leads to directly use raw forcing data from both Canopy and Met towers. Applying the influence of the emergent tree, other CLMML tests have 50m forcing height induced by 33 m

wind speed and have 43.7 m maximum canopy height, with applying different LAD distributions [Figure 5.3] and different wind profile models.

5.2.4 Simulation Setup and Analysis

For CLMML, 30 vertical nodes were used in this study, and the original CLMML code (CLM-ml v0) (Bonan et al., 2018) was modified for the study purposes and inserted in CLM5 code at the middle of the running flow ("clm_driver.F90") to execute CLMML. It is because the update of CLMML is only for canopy process. CLMML runs independently, but it needs other required modules (e.g., soil temperature module, input/output module, etc.) and default variables (Bonan et al., 2018).

Various simulation settings are assigned for the inter-comparison between single canopy models (CLM4.5/5), multi-layered CLM (CLMML), and observation. First, it requires us to add additional equations such as CO_2 profile scheme because CLMML does not calculate it yet. Then, other modification and improvement are mainly to address this complex terrain, which affects leaf distribution, wind/turbulent magnitude, etc. Most of the settings for the all simulations are the same as the default using the satellite phenology (SP) mode. As discussed in previous study, the default parameter is not modified if there is no significant improvement (Song et al., 2019). The major differences from the previous study are the model change (CLMML) for canopy process and forcing height. The additional simulation setup and analysis as below.

This study site presented three main challenges: (1) how to setup the wind forcing data, (2) how to address the complicated canopy shape in the model, and (3) which height the eddy-covariance (EC) measurements actually represents over this sloped surface.

First, applying wind speed data from the canopy gap (EC 33 m) on the Canopy tower as the forcing can be better than on using the data from Met Tower for model simulation. Normally, forcing data for the land surface model is measured on a flat surface like as Met Tower. For example, wind data measured at 10m height over grassland can be applied as forcing value, because it 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 5.2]. Therefore, wind data from the two places were assumed replaceable for each other, and they were applied as forcing data in the previous study for CLM4.5/CLM5 (Song et al., 2019), because the two are highly correlated. However, using such forcing data did not give reasonable simulation results as shown in the previous study. Even a wind speed value which is around two times higher (estimated from the Met tower using a logarithmic relationship) did not give better results. It is possibly because the differences in the turbulent process for flat grassland and sloped canopy surface are not fully understood, which makes it difficult to connect the two locations. Given this result, another idea is tested here to see if the wind speed at the Canopy tower can be used to track a whole wind profile, considering the influence of emergent tree. The basic idea is to use the wind data as forcing to track upward and downward from the middle height of the two canopies, not for downward tracking only.

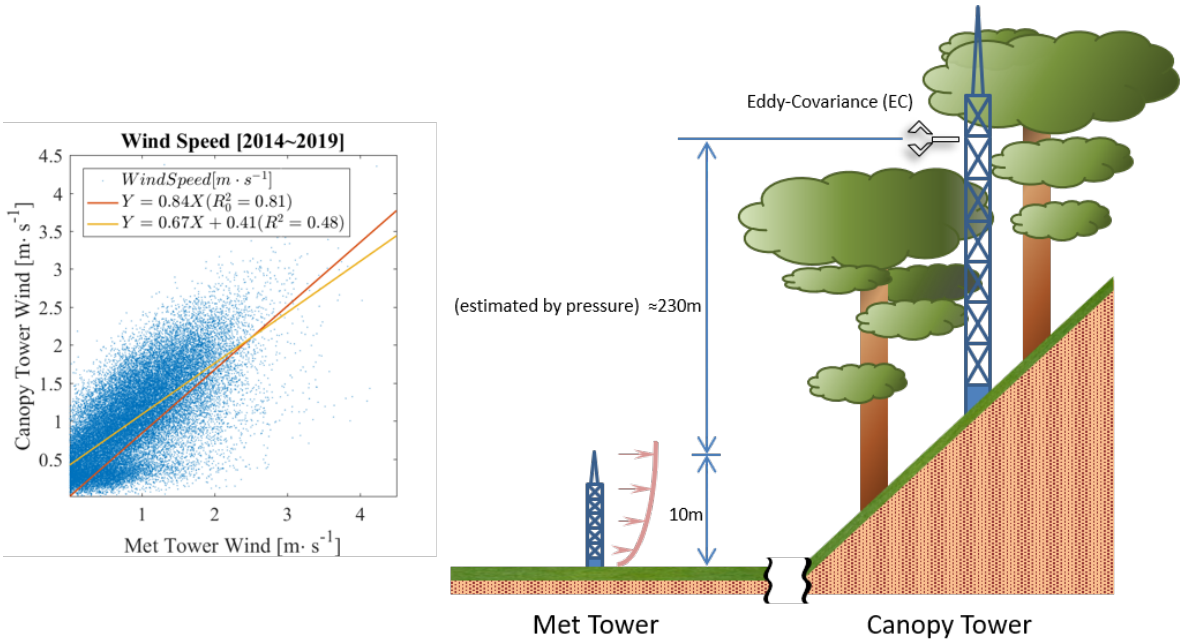

Figure 5.2: Simplified illustration of two different wind sensors. R_0^2 is R-squared without intercept. CLM only takes wind value higher than $1\,\mathrm{m}^2 \cdot \mathrm{s}^{-1}$ to avoid numerical error (Lawrence et al., 2018)

Additionally, the previous wind profile model (RSL) does not perfectly reflect LAD distri-

bution below the maximum canopy height H_{max} (See Appendix B.2.2). Therefore, we need to investigate whether a wind profile model responding to different LAD distributions can result in different model performance. On the other hand, the RSL model is a practical and reliable method. RSL is the updated version of MOST to reflect canopy effects, incorporating the stability theory which has been widely used and proved through many studies. A mixed-model allows for both advantages; here, RSL is applied above the maximum canopy height and the first order closure model is numerically solved in lower portions. Additionally, the computational time is minimized in this test by applying wind profile shapes from the previous step as the initial condition.

The displacement height in RSL is derived based on the centroid drag on the canopy (Harman and Finnigan, 2007). Therefore, the displacement heights for a sloped surface could be relatively lower than those for a flat area, because the maximum canopy height H_{max} and the spatially-averaged height H are not the same [Figure 5.1]. The sloped place produces two parameters: single (flat/average) canopy height H and the maximum canopy height H_{max} especially if all tree's canopies are homogeneously overlapping within a horizontal grid. Then, the canopy height for displacement height (d) may be estimated from average height H, not from H_{max}, if the centroid drag is considered through this sloped canopy.

Secondly, the influence of LAD distributions, maximum canopy height, and displacement heights on this study site (Canopy tower) were assessed. A fitted LAD via measurements is defined as the true LAD profile here. The actual LAD distribution is determined with light sensors and light extinction model, and the shape shows two major canopies appearing in the profile as expected (Andrews et al., 2019). Another possible shape was found in previous literature, where Bonan et al. (2018) describe LAD distributions for deciduous trees and pine trees based on the Beta distribution. We found that pine trees (p = 11.5 and q = 3.5) were most similar to our observed LAD distribution: a tall tree with dense canopy at the top. Canopy height in CLM4.5/5 is 35m as a default value on a flat surface. On the other hand, in this study site, through dual canopy profile model as below, the fitted value for average canopy height is H=39 m and maximum canopy height is about H_{max}=43.7 m caused by topographical slope effect. The dual-canopy model is a

mixed distribution based on a single-canopy model. The Beta distribution function is used for the single-canopy model as in (Bonan et al., 2018)

$$f_{LAD,1}(z,H) = \frac{L}{H} \cdot f_{Beta}(z/H,p,q) + \frac{S}{H} \tag{5.1}$$

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 the Beta distribution. Then, the dual-canopy LAD model can be written using mixed-distribution as

$$\begin{aligned} f_{LAD,2}(z,H_d,r,\Delta H) = r \cdot f_{LAD,1}(z+(r-1)\cdot\Delta H, H_d) + \\ (1-r)\cdot f_{LAD,1}(z+r\cdot\Delta H, H_d) \end{aligned} \tag{5.2}$$

where $f_{LAD,2}$ [$m^2 \cdot m^{-3}$] is the dual-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 5.1]. 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 about LAD model may be found in Appendix B.2.4.

Finally, the eddy-covariance system is located above the lower (downslope) canopy but still between the two major trees, making it difficult to identify what the flux measurements actually represent. This is because the surface is not flat, allowing possible interference from a nearby emergent tree upslope. Here, measured data were compared to simulation results to assess three possible hypotheses using the aid of the multi-layer model. carbon and water flux can have relatively large scale compared to the turbulence scheme. The wind and its turbulence can be regarded as the forcing to transport species (more fit to Eulerian approach). On the other hand, the carbon and water vapor are the actual species which can move long distance (close to Lagrangian ap-

proach). 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 has a footprint which depends on wind direction and the measurement height (Burba, 2013). Moreover, tracking the source can be complicated in the steep area because the footprint can be small for upslope and large for downslope. Wind direction does not normally follow an inclined plane in this study site which makes possible to assume that the upslope and downslope footprints can be counterbalanced. However, this approach does have uncertainty.

We can look the EC data in three different ways: (H0) EC can represent the full flux of the area as in the previous study: EC versus total fluxes from CLM, (H1) EC measures the partial flux of the area which does not contain the residual flux from the emergent tree above the EC measurement at 33 m: EC versus fluxes at 33 m from CLM, and (H2) EC represents a mixture between the top and EC height at the Canopy tower. This test is possible due to a multi-layered model. If EC flux was not falling between Top (H0) and 33 m flux (H1), then we can conclude that model excessively over or under-estimates the fluxes. If H2 is acceptable, then, we can conclude the overestimation of the CLM caused by the error photosynthesis parameter like as previous study, or/and the underestimation caused by low incoming solar radiation at lower canopy (light profile error) which would indicate that light-extinction model may be too simplified for sloped canopy and possibly for sparse canopy having various heights.

All simulation results and observations were compared using the vertical profile of into daytime versus nighttime and wet days versus dry day. Upscaling (or spatially normalizing) 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 dirunal variations about the two variables were estimated as:

$$X = \left(\sum_z LSAI_z \cdot x_z \right) / \sum_z LSAI_z \tag{5.3}$$

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

Finally, the simulation setting is listed [Table 5.1], 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 gives different canopy shapes in simulations, which affects both fluxes and other meteorological variables. R_{10} and Q_{10} are parameters for soil respiration. For the wind model, MOST and RSL are different schemes and their difference only affects above the canopy ($\geq H_{max}$) in this study. We use the term (Numerical) for when a turbulent model for under the H_{max} is changed from RSL (closed-form) to the First-order closure model. The LAD distribution was described previously but the soil and the wind model will be discussed in later section. The different canopy shapes are plotted in [Figure 5.3].

Table 5.1: Model list and the name is abbreviation

#	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

5.2.5 Additional Calculations

In addition to introducing new LAD distributions, we added or reformulated three aspects of CLMML: reformulated a wind speed profile scheme, and added a CO_2 profile scheme and a soil respiration flux. The equations are discussed briefly here; more detailed information may be found

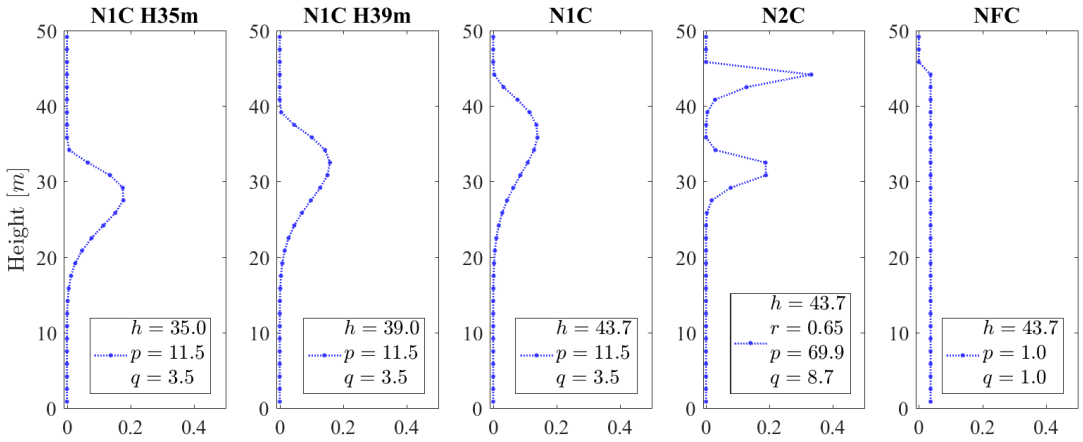

Figure 5.3: Various LAD distributions by the two parameters (p, q, r) and heights $(h = H_{max})$ for this study [Table 5.1]. 30 nodes are used here.

in Appendix B.2.

A first-order closure model of wind profile for the eddy-diffusivity is used in this study to reflect different LAD distributions because RSL (closed-form but approximation) cannot fully reflect LAD distributions as mentioned previously. The model follows previous work (Launiainen et al., 2011; Katul et al., 2004; Drewry et al., 2010) as

$$K_m \frac{\partial^2 u}{\partial z^2} + \frac{\partial K_m}{\partial z}\frac{\partial u}{\partial z} - C_d a(z) u^2 = 0 \tag{5.4}$$

where K_m is the eddy diffusivity for momentum, u is wind speed, z is height, C_d is drag coefficient (0.25), and $a(z)$ is LAD (Launiainen et al., 2011). The eddy diffusivity for momentum is

$$K_m = l_m^2 \left| \frac{\partial u}{\partial z} \right| \tag{5.5}$$

where l is mixing length given as

$$l_m = \begin{cases} k_v z, & z < \alpha' H_{max}/k_v \\ \alpha' H_{max}, & \alpha' H_{max}/k_v \leq z < H_{max} \\ k_v(z-d), & H_{max} \leq z \end{cases} \tag{5.6}$$

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 $0.667 \cdot H_{max}$ but it varies in CLMML (Bonan et al., 2018) [See Appendix B.2.1]. For this study, the displacement height is estimated based on the average canopy height H or H_d, so the height is calculated based on the concept of dual-canopy model ($f_{LAD,2}$) to apply the slope effect.

In different settings as [Table 5.1], the wind profile is mostly (except #1-4) estimated based on EC data (u_{EC}) placed at the middle of the profile system [Figure 5.1; Figure 5.2]. For the first-order closure model, the ordinary differential equation (ODE) is solved based on two wind speed values as the boundary conditions which are at the EC location and the ground (u=0). The wind speed is calculated for the higher level until the maximum canopy height, using the wind speed and its vertical gradient at the EC location. On the other hand, this calculation becomes straightforward in the RSL model because it has an analytical solution in which the key function depends on the height ($u(z)$=$f(z)$) [See Appendix B.2.2]. Wind speed as a function of height is estimated using the reference wind speed (u_{EC}) at the EC location (z_{EC}) as

$$u(z) = u_{EC}(z_{EC}) \cdot f(z_{EC})/f(z) \tag{5.7}$$

The method to determine CO_2 concentrations as a function of height (i.e., CO_2 profile) is similar to other profile estimations like as water vapor in CLMML, described by Bonan et al.

(2018):

$$\rho_m \frac{\partial C}{\partial t} - \frac{\partial}{\partial z}\left(\rho_m K_c(z) \frac{\partial C}{\partial z}\right) = \left[f_{c,sun}(z) f_{sun}(z) + f_{c,sha}(z) f_{sha}(z)\right] a(z) \tag{5.8}$$

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 and it was assumed the same as K_m in this study based on the previous study that the ratio appears 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. Soil respiration R ($\mu mol{\cdot}m^{-2}s^{-1}$) is also applied in CLMML as a source term. The expression follows as in Launiainen et al. (2011) which has

$$R = R_{10} Q_{10}^{(T_g - 10)/10} \tag{5.9}$$

where R_{10} and Q_{10} are parameters which have 0.3 and 2 ($R{\approx}0.7$ at 22 °C) (Launiainen et al., 2011), and T_g is ground temperature (°C). These parameters are also fitted using measured data in this site which have 2.4 and 1.7 [Figure B.2a in Appendix B.1]. This fitted soil respiration model (called as 'Q1' in this study) tends to produce very high rates ($R{\approx}4.5$ at 22 °C). Also, including CLM5/CLM4.5, CLMML have a relatively high ground temperature which makes soil respiration rate too high (model's error). Therefore, for 'Q1' simulation the ground temperature is directly predicted by the simple regression model [Figure B.2b in Appendix B.1].

Applying these additional equations and to conduct its inter-comparison study, the CLMML and CLM5 were simulated from 2014 to 2017. Also, CLM4.5 was also simulated because CLMML is based on CLM4.5. However, the comparison is mostly with CLM5 if there is no significant difference.

5.3 Simulation Results and Discussion

Total net radiation showed no significant differences between any single layer model (CLM4.5 and CLM5) and a multi-layer model (CLMML) [Figure 5.4] because it was estimated from the same radiative transfer parameters, particularly albedo. A small daytime difference may originate

due to multi-layered radiative transfer scheme, and more significantly due to the fact that the multi-layer model was able to resolve energy segmentation at each height through the vertical profile, which produces different total net radiations from different LAD shapes [Figure 5.5a; Figure 5.5b; Figure 5.5c]. The impact of in-canopy variation can be detected via flux comparisons. The night-time net radiation has been decreased when the case has a higher canopy height (see net radiation chart in supplementary data). In this case, the wind turbulence model highly affects the energy exchange in the night time. A net radiation profile including the PAR profile in the daytime is greatly influenced by canopy shape [Figure 5.5a; Figure 5.5b], and we need to note that the net radiation profile is also affected by light–extinction model. Consequently, this vertical variation of net-energy affects physiological activities such as GPP and transpiration (TR), which has similar profile shape (shown later).

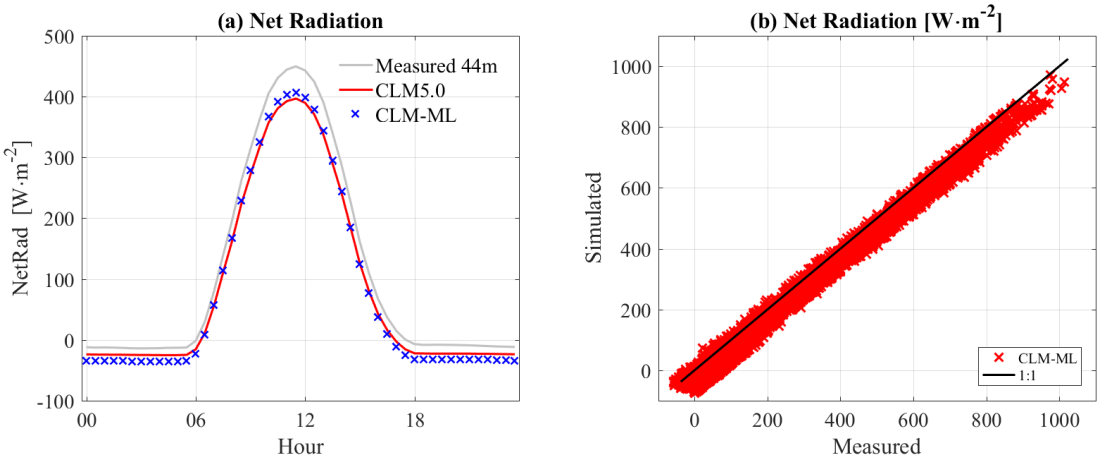

Figure 5.4: (a) and (b) - comparison of net radiation between CLM and measurement on Canopy Tower at 44m.

Simulated net radiation and PAR profiles for the dual-canopy simulation (N2C) have a two-steps stair shape, and we note that contrary to expectations, the amount of received energy is significantly different between the two upslope and downslope trees [Figure 5.5a; Figure 5.5b]. The inconsistency arises because in effect the two canopies spatially overlap slightly [Figure 5.1].

This error likely occurs because the measured LAD profile has a relatively small footprint. On the other hand, the PAR profile follows the observations well, compared to other shapes [Figure 5.5a], and the predicted net radiation profile with the two canopies (N2C) can be more reasonable than with other shapes.

Additionally, the evapotranspiration (ET) fraction profile tends to converge to the 0.7 at a high leaf area density and a high energy exchange [Figure 5.5b; Figure 5.5d]. These profiles indicate that the ET flux is a main contributor to the energy exchange in this site. Also, it shows the importance of a canopy shape because the received energy profile depends on the leaf distribution. From the PAR profile shape, and the profiles of net radiation and sensible heat flux, it seems that the very thin thickness of leaf density is contributing the most of energy exchange including ET. Therefore, an accurate canopy model such as physiologically-based model is very important.

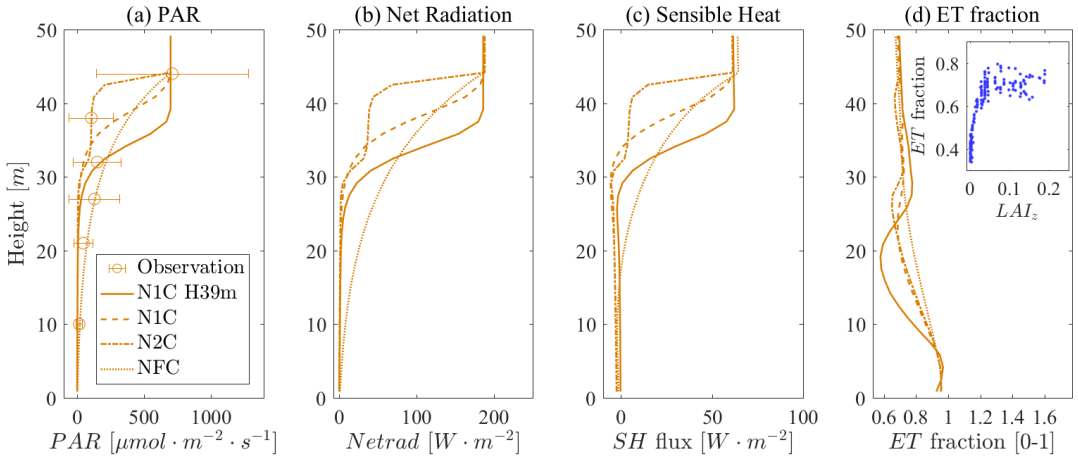

Figure 5.5: (a) PAR profile, (b) Net radiation profile, (c) Sensible heat flux, and (d) Evaporative fraction in daytime. These profiles are simulated through multi-layered CLM (CLMML) with different structures [Table 5.1].

The shape of the wind profile generated by the numerical method tends to follow the observations within the canopy well [Figure 5.6]. Applying the more realistic canopy height and shape (dual-canopy formulation) gives a more plausible wind profile [Figure 5.6c]. Here, the different shape of the canopy profile does not seem to yield large differences in wind speeds (i.e., N1C vs

N2C) [Figure 5.6c]. Some differences in the wind profile can be identified between the numerical estimation (N1C) and RSL (R1C), however, these differences have a little impact on the simulation of final flux at the top which will appear in later comparisons and Supplementary Data. Our analysis indicates that the change of the canopy heights produces a notable impact on the wind profile. The different heights for this test case consist of the mean canopy height 39 m and the actual maximum height 44 m due to the slope effect. Also, the displacement height (d) would be lower than usual [Figure 5.1]. RSL profile with lowered d (R1C) seems to worsen the results than with normal d (R1C H44m), but we need to note that RSL cannot reflect the canopy shape. The first-order closure model seems to have some improvement, given its low d value (see N1C H44m vs. N1C) [Figure 5.6b].

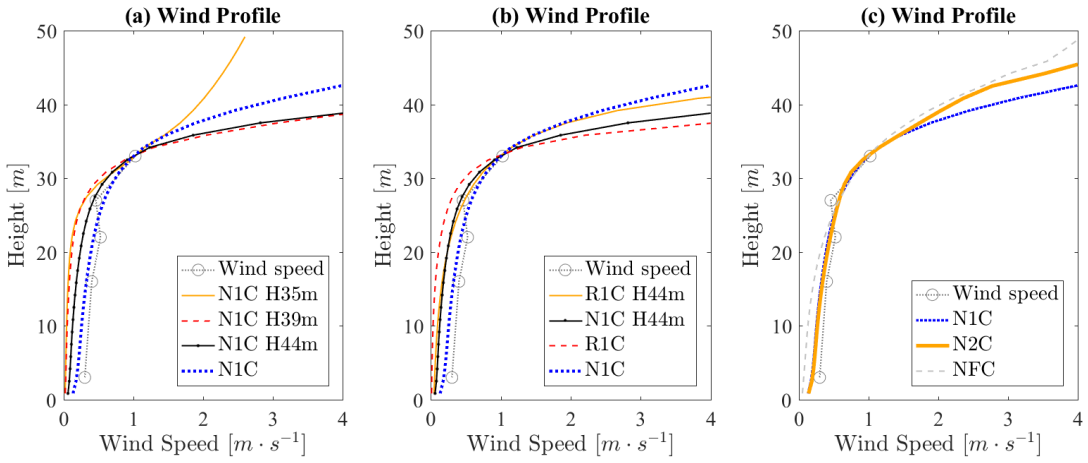

Figure 5.6: Wind profile. (a) Wind profiles with different canopy heights. H indicates H_{max}=$H(d)$ [Table 5.1]. (b) Wind profile with different models. (c) Wind profile with different LAD distributions.

The multi-layered model can give significant improvement to the prediction of leaf wetness [Figure 5.7]. The leaf wetness profile for wet days highly depends on the shape of LAD (see in [Figure B.3] and Supplementary data). Also, the increase in nighttime evaporation has an important role in moderating the over-fluctuation in diel variation compared to the previous study (Song et al., 2019). The overall result indicates that the higher portions of the canopy in CLMML tend to follow

well measured values, but the lower portions of the profile, which also have low leaf areas, tend to hold more water than observations would suggest [Figure B.3 in Appendix B.1]. This is possibly caused by low net radiation at the lower heights as shown in the profile shape of net radiation. We cannot visually identify the impact between two different turbulent transfer schemes (R2C and N2C) but the small difference can be detected in the nighttime and wet day (see supplementary data). Nonetheless, the multi-layer model provides more reasonable results as displayed in the comparison using up-scaled leaf wetness [Figure 5.7]. Considerable measurement error may be expected from these sensors due to their shape and low spatial coverage, but they do provide insight into diurnal variations in wetness.

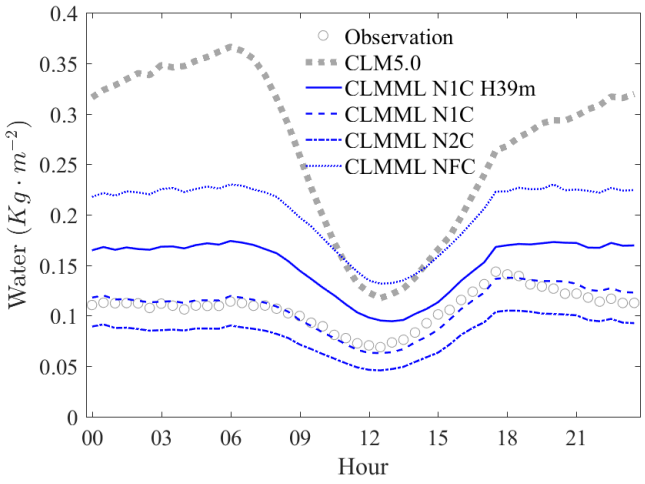

Figure 5.7: Diurnal variation plot for canopy water which is up-scaled from profile results [Figure B.3 in Appendix B.1] via Eq. (5.3). Here the wetness is not used to avoid the confusion by the power term (2/3) in CLM. This canopy water represents the ratio between the current canopy water and the maximum canopy water $W_{max} = 0.1\mathrm{kg}\cdot\mathrm{m}^{-2}$ without the power term, because W_{max} is a constant and total LSAI has only temporal variability. The water on y-axis represents the total water on the canopy : the LSAI is approximately $\approx 6\ \mathrm{m}^2\cdot\mathrm{m}^{-2}$.

Modeled profiles of air temperature, relative humidity, and CO_2 concentration show general agreement with the observations, indicating that these variables were mainly affected by turbulence-transfer parameters, especially by the maximum canopy height H_{max} and LAD distributions [Fig-

ure 5.8; Figure 5.9; Figure 5.10]. The night-time profiles are noticeably improved compared to the single-layer model. Most profiles seem within the standard deviation of observed profiles (see error bar plots), but some cases with low canopy heights such as N1C H39m and N1C H34m tend to have more errors especially in CO_2 profile.

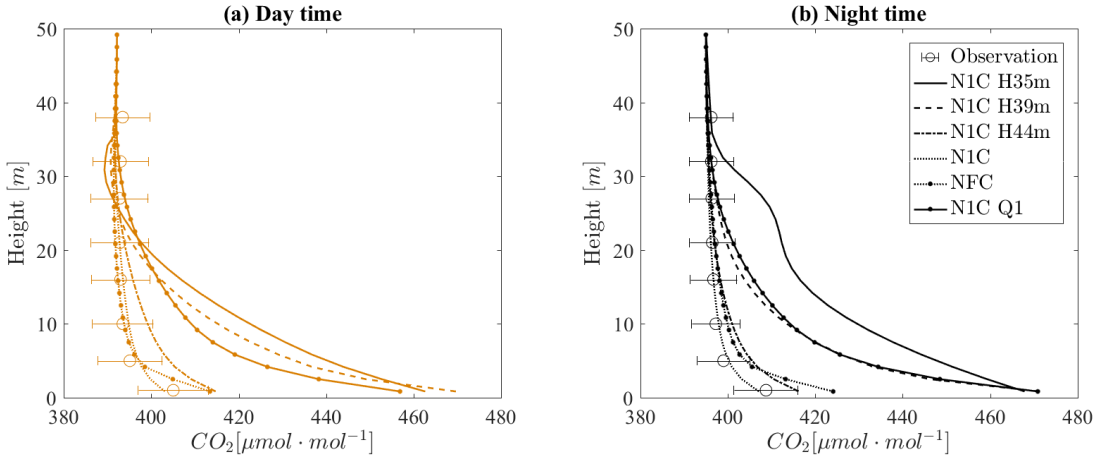

Figure 5.8: CO_2 concentration profile, with observations from AP200 system

Both the CO_2 profile and CO_2 flux were highly influenced by the soil respiration [Figure 5.11; Figure 5.8]. Applying parameters from Launiainen et al. (2011) resulted in reasonable predictions but applying our measured soil flux (N1C Q1) made the comparison deviate from expectations. However, it is not yet certain which one is correct, due to the significant error of GPP prediction in CLMML that still exists.

For CO_2 and H_2O fluxes, both H0 and H1 types of interpretation do not yield a successful comparison between EC data at 33 m and simulated fluxes, but the H2 type seems 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, the multi-layer model seems to overestimate the total flux in the daytime for all the simulated cases [Table 5.1] [Figure 5.11]. If we assume H1 is correct, the simulated fluxes are mostly underestimated. There are some exceptions such as a simulation (N1C H35) having 35m canopy height which overestimates CO_2 flux but it did not match the

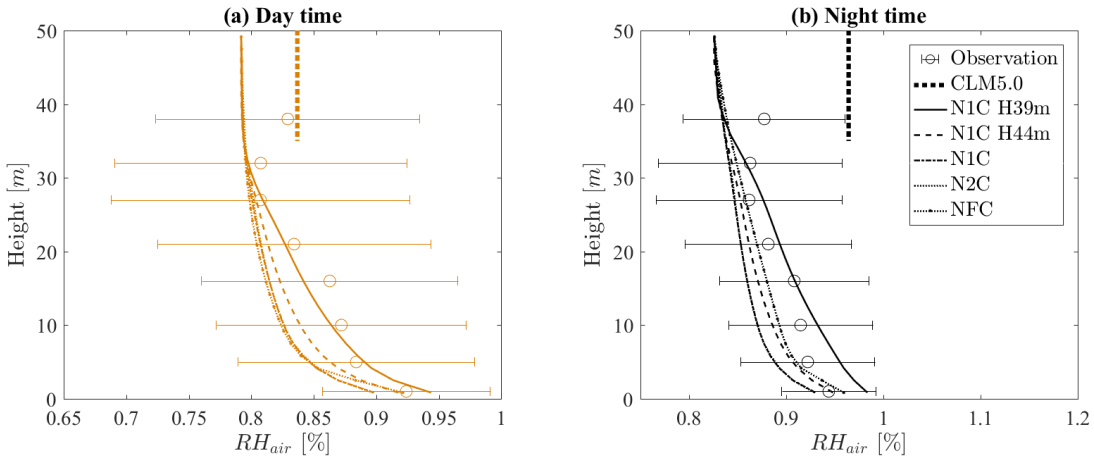

Figure 5.9: Relative Humidity (RH) profiles with the profiled observation.

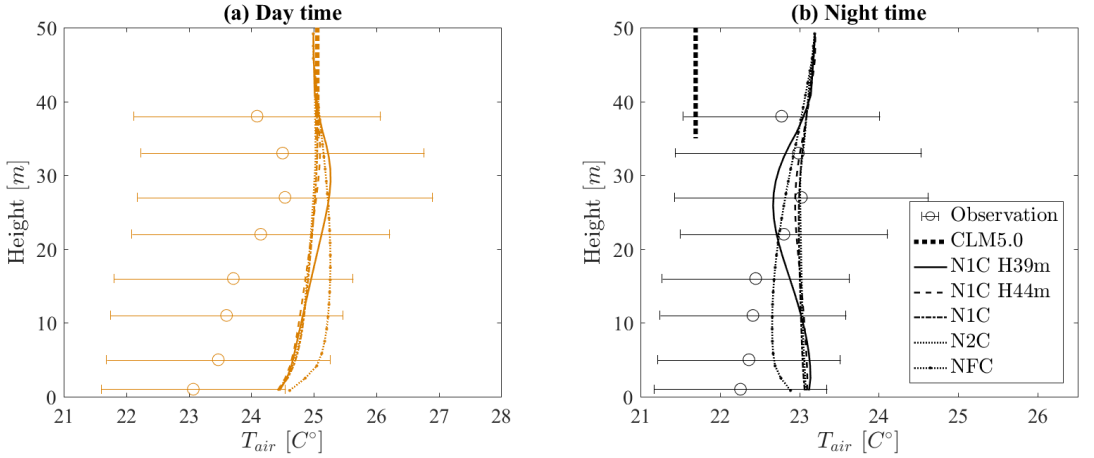

Figure 5.10: Air Temperature profiles.

CO_2 profile either [Figure 5.8]. The other 39 m simulation (N1C H39) and the 44 m flat canopy simulation (NFC) matched observed H_2O fluxes at 33 m but their wind profiles and leaf wetness are not well predicted. Therefore, both hypotheses were not supported and the 33 m observation (eddy-covariance) may represent a height between the two places (H2 case). These results are also supported via transpiration (TR) observations, where total sap flow rates is higher than 33 m eddy-covariance and lower than the simulated TR rate (Aparecido et al., 2016).

These simulated carbon/water fluxes as well as their concentrations indicate that as compared

with the single layer model, sub-canopy models produce significantly different results and are sensitive to changes in structural parameters. Particularly, canopy parameters that affect the turbulence model and light penetration alter the resulting understory concentrations of trace gasses. For instance, GPP and transpiration (TR) predictions are mainly influenced by radiative transfer [Figure 5.13], which is mainly affected by LAD distribution. Fluxes CO_2/H_2O and their concentrations are more altered by changing turbulence-related formulations or parameters (e.g., canopy height, LAD distribution) [Figure 5.11; Figure 5.12] (also see supplementary charts). These intrinsic variations, caused particularly by canopy shape, produce diverse fluxes of land-atmosphere interaction. Here, we need to notice that CO_2/H_2O fluxes, their concentrations, and GPP/Transpiration are different. These fluxes represent final fluxes toward the atmosphere/in-canopy air which are affected by the turbulence model. GPP/TR indicates source and sink through physiological activity.

The multi-layered model tended to show higher night time evaporation rates, which is caused by improved turbulence scheme by model itself (RSL) including model parameters and different LAD. These changes consequently contribute to the improvement of leaf wetness compared to the single-layered model. The nighttime water flux mostly occurs at the dense top canopy [Figure 5.12]. However, the comparison problem for the water flux exists similarly to CO_2 flux when compared with EC data, which makes diagnosis difficult. Through the aid of concentration profiles, the in-canopy behavior of the model can be diagnosed. For more accurate partitioning the water flux, all models related to canopy water need to be re-investigated with more observations to capture the effects particularly produced by the topographical complexity.

The profile of TR rate reasonably mimics observation but the predicted total flux (TranDom) is still overestimated [Figure 5.13]. This result supports the idea of the previous study (Song et al., 2019) which highlights the possible issues with light-limited photosynthesis models. In this site, the sap-flow data can regraded as profile data, as it is measured from trees with a range of canopy heights. The data has three different categories: Sub-story, Mid-story, and Dominant canopy are separately measured, and these are accumulated toward the top canopy for the profiled plot. The heights are estimated as $H_{sap} = H + \Delta H(1 - r)$ according to the previous sap-flow study which

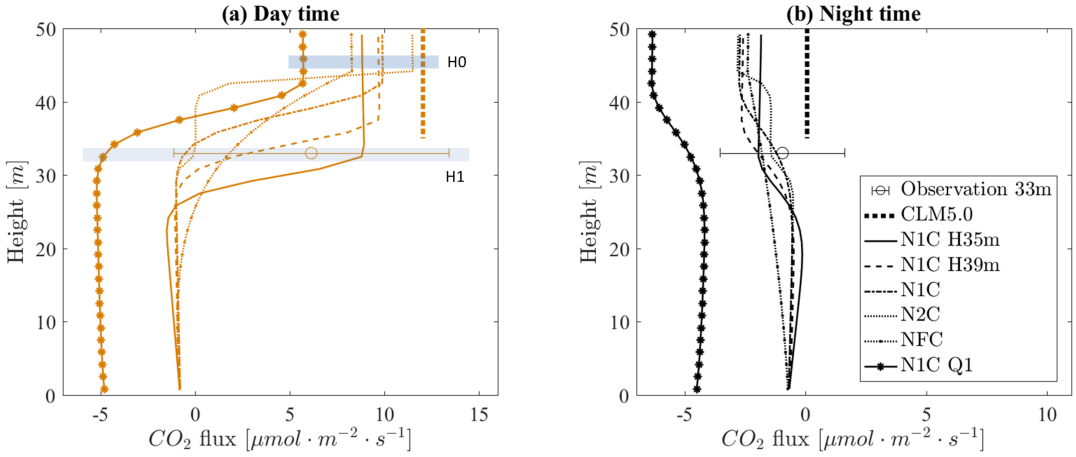

Figure 5.11: CO_2 flux profiles simulated using under different canopy parameters [Table 5.1]. For the observation from EC measurement, the circle is the mean value and the bar indicates standard deviation.

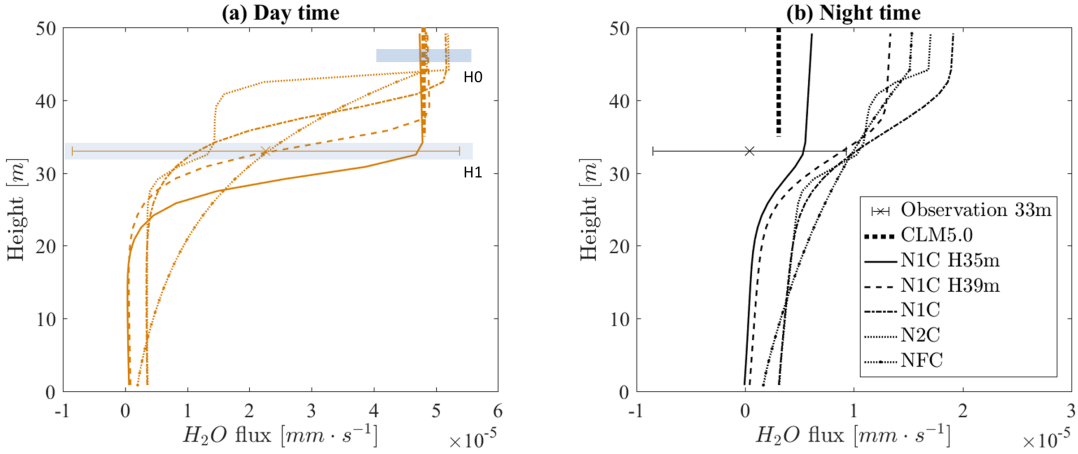

Figure 5.12: H_2O flux profiles in the same way of [Figure 5.11].

gives possible each maximum canopy height for flat surface as H = 11m, 27m, and 39m. Other parameters for sloped surface are ΔH = 13.5m and r = 0.65, which result in 15.725m, 31.725m, and 43.725m for the Sub-story, Mid-story, and Dominant canopy.

Night time improvements due to the multi-layered scheme can be also detected in the diurnal variation of leaf temperature [Figure 5.14]. The overall temperature is increased but the amplitude of leaf temperature, as well as air temperature [Figure 5.10], is reduced and the cycle follows well

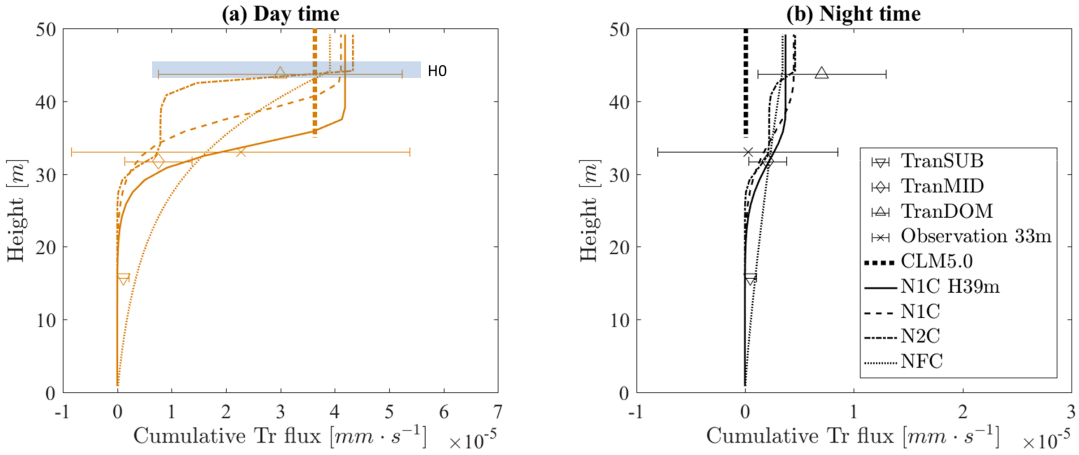

Figure 5.13: Transpiration (TR) profile in the same way of [Figure 5.11]. It should be noted that the Observation 33 m is EC which represents total H_2O flux including TR. TranSUM, TranMID, and TranDOM are from sap-flow data, which represent in 15.725 m, 31.725 m, and 43.725 m for the Sub-story, Mid-story, and Dominant canopy

the observation (see N1C case). Through the multi-layered model, other variables such as RH and Air Temperature were also improved compared with the single-layer model (CLM5) [Figure 5.9; Figure 5.10].

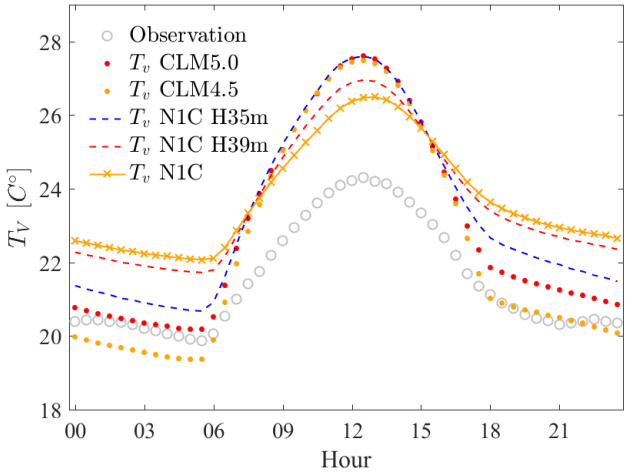

Figure 5.14: Diurnal plot of Leaf Temperature (T_V) via Eq. (5.3).

5.4 Summary

In this study, we tested multi-layered CLM (CLMML) (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 on a complex terrain. Here, additional calculations were included in the model to test different turbulent transfer models (i.e., wind profiles) and various canopy shapes, as reflected in leaf area density variations with height [Figure 5.3]. Additionally, a new formulation for predicting in-canopy CO_2 concentrations, which includes the effects of soil respiration, was also estimated through CLMML to compare with our available data set.

Overall, the top-of-canopy energy balances predicted by CLMML were similar to those from CLM due to the same radiative transfer model. However, CLMML's main advantage was to reproduce trace gas concentrations, and micrometeorological variables with the subcanopy, allowing for paritioning of trace gas fluxes as a function of height in the under-story. The small alteration of the net radiation comes from a different energy partitioning caused by the change of model structure (e.g., multi-layer and RSL). The simulated meteorological profiles such as air temperature, RH and CO_2 concentration are not perfectly matched with the observations. However, the mean profile falls within the standard deviation of observation in most of test case. We also found the CO_2 concentration profile and its fluxe are very sensitive to the soil respiration, which indicates we need to focus on soil activity for tropical site.

With the aid of multi-layer scheme, this study showed it is possible to model a complex terrain by including its atypical features. The general wind speed model cannot be simply applied in a mountain area (sloped surface) to accomplish comparisons because the surface is not flat and accordingly overall LAD distribution becomes a mixed-distribution (dual-canopy) [Figure 5.1]. Also, the displacement height is derived based on the centroid drag on the canopy area, which results in the lower displacement heights by the slope effect.

Furthermore, the RSL model embedded in previous CLMML for under canopy may not be adequate to predict a turbulent transfer because it cannot reflect LAD distribution. Therefore, a

first-order closure scheme Eq. (5.4) is applied to explore its effect. The wind forcing data located at the middle height of the canopy profile was successfully included in the model [Figure 5.6]. This shows the possibility to overcome the difficulty of data use/ modeling at abnormal site. It also highlights that normal forcing data (such as Met Tower) cannot be easily applied for such complex terrain, because the wind profile was abruptly increased at the higher location [Figure 5.6].

Applying different types of wind speed models and parameters clearly affects the wind profile shapes, and the canopy height is most influential parameter for overall performance. However, changes between numerical and RSL methods and different displacement heights have a similar performance for the other variables (e.g., temperatures, etc.), especially with single-canopy profile. The small difference can be detected for leaf wetness and H_2O flux when dual-canopy LAD is applied (see supplementary data). This similarity can be mainly because RSL model is derived based on the first-order closure model. In contrast, some small differences are identified here (see actual value in supplementary data), which indicates there is still the possibility of having a stronger influence in a different environment. Since we used only several leaf profiles, fixed LAI, etc., different ambient conditions can show different outcomes. More investigation is necessary.

This study highlights that the two schemes, turbulent model and LAD distributions including radiative transfer model, are important for the performance of land surface models. LAD distributions affect the direct/instant energy exchange over the leaf profiles, so a different LAD distribution produces significantly different outcomes for trace gas and momentum fluxes as well as state variables(e.g., CO_2/H_2O flux, GPP, TR, leaf wetness, and leaf temperature). These differences are directly attributable to the influence of light penetration through the canopy LAD profile. They are also influenced by the effect of wind profile (turbulent transfer) due to in-canopy feature (canopy height, LAD distribution). Between the two schemes, the micrometeorological profiles in the air (e.g., air temperature, RH, and CO_2 concentration) are not solely/directly affected by LAD distributions (e.g., bell shape) unless the canopy shape is too different, such as a flat canopy (NFC). They are rather more affected by parameters for the turbulent model such as the canopy height.

This study emphasizes the importance of layered structures in capturing the behavior complex

surface systems and supports previous multilayered model studies (Jiménez et al., 2011; Ogée et al., 2003; Pitman et al., 2009; Bonan et al., 2014). Through this test, we found that updating the model structure from the single-layer model (CLM4.5/CLM5) into the multi-layer model (CLMML) results in improvement of water and energy partitioning, by allowing the exchanges of matters and energies between vertical layers. Moreover, it makes further improvement possible by applying a more realistic canopy shape, which affects radiative transfer and turbulent transfer. This multi-layered model loosens the limitation of simplified model and allows more variability caused by the internal structures (below canopy) (Ryder et al., 2016). For instance, the vertical segmentation of vapor transfer as a function of height gives a notable improvement in leaf wetness predictions. Also, GPP and TR show that different canopy structure results in different total fluxes through vertical distribution of sources and sinks. Simulated total GPP (by vertical summation) is increased but the top flux toward the atmosphere is reduced by a multi-layered turbulent scheme and by the storage terms. These features cannot be captured using a single-layer model. Updating the MOST model (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.

Perhaps more importantly, the profile measurement system provided additional, more detailed information to diagnose the model performance than is normally possible through the single-layer model. For instance, the representative height for the eddy-covariance (EC) in this site was approximately expected as H2 case through the analysis of the profiled results using CLMML. In this analysis, some cases seem well matched at a certain height between EC and CLMML, but they fail to reproduce the associated variables (e.g., N1C H39 seems to well estimate H_2O flux at 33 m but its wind profiles, leaf wetness, and CO_2 profiles are not well predicted). The profile results related to TR show comparison between sap-flow data and simulated TR has good agreement but still overestimate photosynthesis. This result supports the previous idea that the parameter for the light-limit photosynthesis needs to be updated.

5.5 Conclusion

In conclusion, the multi-layered CLM have intrinsic improvements by reflecting complicated structures under the canopy compared to the single-layered CLM. This study shows possible potential for the further improvement of LSM using this multi-layer scheme by the more detailed interpretation of the natural phenomenon such as updating wind turbulent scheme and leaf area profile. Moreover, the vertical profile scheme helps to examine model's ability to capture the in-canopy processes resulting in the final land-atmosphere interaction. It helps to track the error of sub-model in detail by comparing with such spacially rich measurement.

However, the multi-layered CLM still cannot reflect topographical complexity. Although this study attempts to update the original model to overcome such a deficiency, the model fails to sufficiently capture slope effects or to find a clear answer for model improvement. An improved capturing light-penetration scheme, especially reflecting a variety of combinations between the surface slope and solar location, may be the first task to solve. In additionally, more observations are necessary, particularly data on light profiles but also wind speed profiles. To fully understand such complex mountain forest, vertical profile data are necessary but also the data reflecting horizontal spatial variability are also required, which may be available at only few sites.

6. SUMMARY

The terrestrial water and carbon cycles are important in the Earth system which influences both living things and their climate. Water migration transverses multiple sub-systems in the Earth (e.g., vegetation, soil, groundwater, river flow, precipitation). Therefore, their internal behaviors and interactions with the atmosphere or with the other sub-systems need to understand, to predict future water availability but also the climate impact. Particularly, we need to capture the relatively fast responsive process such as the land-atmosphere interaction which highly affects future climate.

Particularly, soil evaporation and vegetation evapotranspiration including photosynthesis are key processes in land-atmosphere interaction. However, their mechanisms are not fully understood yet. Therefore, it is important to keep studying the soil and vegetation processes for the better understanding and accurate LSM, through model diagnose and update.

In Section 2, a new relationship for the saturated front depth (also known as the characteristic length) was derived under the condition of a shallow water table. Previously derived models have considered only single-directional flow to determine the front depth, and the comparison tests resulted in significant disagreement with a mass-conservation model (Richards Equation) with respect to the front depth and the total mass loss. Through this study (Section 2), a two-directional approach (vertical and horizontal) was derived to capture both the model improvement and mass-conservation. The relationships were derived and proved using two different equations, the anisotropic Darcy's law and the Richards equation.

In this study (Section 2), the test of the two-directional approach results revealed that the new concept provides better performance in both the front depth and the total mass loss, and this study shows the advantage of analytical/semi-analytical solution against a numerical method. It also implies that with further study it may be possible to predict the front depth depending on the different surface water content.

In the next section (Section 3), relative evaporation models were explored in more detail, and they were reformulated to reflect heterogeneous pore sizes at soil surface based on capillary theory

(Young-Laplace equation) and pore-size distribution. The evaporation models in this study were based on an idea that drying rates are controlled by the two main resistances from an inside soil layer (SL) and a viscous boundary layer (VBL). Within the VBL, some diffusion-based models (e.g., S&M, SCH), which normally consider homogeneous soil pores, have been often used to predict an evaporation rate. Such two models (S&M, SCH) were regarded as having similar behavior each other but they were not sufficiently identical. Also, to apply the heterogeneity of soil surface at VBL, effective pore sizes and effective water contents facing the atmosphere were derived and applied into a diffusion-based evaporation model. Then, the influence of soil-pore heterogeneity was tested accounting for both soil and VBL resistance schemes in the evaporation model.

In Section 3, our investigation found that the two major diffusion-based models (e.g., S&M, SCH) are sufficiently close to each other after the re-interpretation of the model concept. Then, with SCH, the simulation test for the heterogeneous-pore-size model indicates the drying rates tend to be lower than the original models (homogeneous). The reduction was especially notable between the beginning of evaporation (fully saturated condition) and transition point (stage change), which accounts for a large amount of water movement even with small change. Also, this study shows that effective water content tends to have a more significant role than effective pore size in the mechanism of evaporation.

Section 4 and Section 5 examined the performance of the Community Land Models (CLM4.5, CLM5, CLM-ml) against tower and ground measurements from a tropical montane rainforest in Costa Rica. The study site receives over 4,000 mm of mean annual precipitation and has high levels of relative humidity. The measurement tower is equipped with an eddy-covariance and vertical profile systems able to capture various micrometeorological variables, particularly for wet and complex terrain.

In Section 4, results from point-scale simulation for both CLM4.5 and its updated version (CLM5) were compared to observed canopy flux and micro-meteorological data. Both models failed to capture the effects of frequent rainfall events and mountainous topography on the variables of interest (temperatures, leaf wetness, and fluxes). Overall, CLM5 alleviates some errors

in CLM4.5 but CLM5 still cannot precisely simulate a number of canopy processes in this forest. Soil, air, and canopy temperatures, as well as leaf wetness, are still too sensitive to incoming solar radiation rates despite updates. As a result, daytime vapor flux and carbon flux are overestimated, and all temperature differences between day and night are higher than the observations. Slope effects appear in the average diurnal variations of surface albedo and carbon flux from measured data, but CLM5 cannot reflect these features. This study suggests that both CLM models still require further improvements concerning model's structures, but also energy partitioning processes, such as leaf wetness process, photosynthesis model, and aerodynamic resistance model for wet and/or mountainous regions.

Section 5 utilized multi-layered Community Land Model (CLM-ml, CLMML) to test. Some difficulties of applying forcing data and model's evaluations due to topographical complexity overcome through this multi-layered scheme. A wind profile for reliable simulation was traced from available wind data at the middle of the canopy rather than very high height. The wind profile, estimated by the analytical and numerical scheme, and different leaf area shapes were applied in the model below the top canopy height for structure tests. For additional comparison, CO_2 profile model and soil respiration model were also added in this model.

Study results presented in Section 5 indicate that multi-layer CLM has the ability to mimic the under-canopy shapes of micrometeorological profiles of multiple variables (humidity, CO_2, temperature, and wind speed). Especially, the over-amplitude of diurnal variations for leaf temperature and leaf wetness are significantly improved compared to the single-layer model (CLM4.5 and CLM5) in Section 4: the amplitude between the maximum and the minimum value in diel variation are decreased by 67% and 47% with the CLM-ml. It also highlights that sub-canopy structure such as canopy shapes and the parameter for the turbulent transfer model can play a significant role in model performance. More importantly, such a profiled system provides detailed information to diagnose the model performance that is not normally possible through the single-layer model. This study indicates that representing sub-canopy activity through structure improvement can produce more accurate estimations of energy and hydrological processes in CLM. However, more extensive

monitoring of sub-canopies is necessary for further study and to increase model reliability for other sites with complex terrain and irregular vegetation roughness.

Based on this research, the integrated and extended studies are possible for future study. The combination of front-depth theory and a heterogeneous-pore-size evaporation model can be a promising example of further study to catch soil evaporation process via considering both above and below soil surface. Furthermore, this new soil-evaporation model can be applied to the global scale LSM, as can the point-scale model developed for the forest.

Overall, extensive work to diagnose and improve soil and canopy process yielded better representation of the land surface process for advanced LSM. This study highlighted the current model' errors and provided suggestions or improvements. However, an increase in understanding each main system (e.g., soil and canopy) still remains a challenge, as do the computational requirements of highly integrated or fine-scale models. Interpretation of complex nature system and also finding a practical method are necessary, through the model's development and various observations.

REFERENCES

Ali, A. A. (2016). A global scale mechanistic model of photosynthetic capacity (LUNA V1. 0). *Geoscientific Model Development*, 9(2):587.

Aluwihare, S. and Watanabe, K. (2003). Measurement of evaporation on bare soil and estimating surface resistance. *Journal of Environmental Engineering*, 129(12):1157–1168.

Andrews, R., Song, J., Miller, G. R., Cahill, A. T., W., G., and Moore (2019). Micrometerological profiles in a lower montane tropical forest: Variations induced by diurnal cycle, leaf wetness, and seasonality. In preparation for submission to Agricultural and Forest Meteorology.

Aparecido, L. M. T., Cahill, A. T., Miller, G. R., and Moore, G. W. (2017). Leaf surface traits and water storage retention affect photosynthetic responses to leaf surface wetness among wet tropical forest and semiarid savanna plants. *Tree Physiology*, 37(10):1285–1300.

Aparecido, L. M. T., Miller, G. R., Cahill, A. T., and 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.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S. (2001). 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.

Ball, J. T., Woodrow, I. E., and Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In *Progress in photosynthesis research*, pages 221–224. Springer.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain,

M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D. (2010). Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, 329(5993):834–838.

Bonan, G., Williams, M., Fisher, R., and Oleson, K. (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.

Bonan, G. B., editor (1996). *A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and user's guide*. NCAR Technical Note NCAR/TN-417+STR, National Center for Atmospheric Research, Boulder, CO, 150 pp.

Bonan, G. B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science*, 320(5882):1444–1449.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., and 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).

Bonan, G. B., Oleson, K. W., Fisher, R. A., Lasslop, G., and 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).

Bonan, G. B., Patton, E. G., Harman, I. N., Oleson, K. W., Finnigan, J. J., Lu, Y., and 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.

Brooks, R. and Corey, T. (1964). Hydraulic properties of porous media. *Hydrology Papers [Colorado State University]*, 24.

Brutsaert, W. (2005). *Hydrology: an introduction*. Cambridge Univeristy Press, Cambridge, U.K.

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., Knowles, J., and 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.

Buzan, J. R., Oleson, K., and Huber, M. (2015). Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5. *Geoscientific Model Development*, 8(2):151–170.

Cai, X., Yang, Z.-L., Xia, Y., Huang, M., Wei, H., Leung, L. R., and 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.

Campbell, G. S. and Norman, J. M. (2012). *An Introduction to Environmental Biophysics*. Springer New York.

Carsel, R. F. and Parrish, R. S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5):755–769.

Chaney, N. W., Herman, J. D., Ek, M. B., and Wood, E. F. (2016). Deriving global parameter estimates for the Noah land surface model using FLUXNET and machine learning. *Journal of Geophysical Research: Atmospheres*, 121(22):13,218–13,235.

Chanzy, A. and Bruckler, L. (1993). Significance of soil surface moisture with respect to daily bare soil evaporation. *Water Resources Research*, 29(4):1113–1125.

Chen, Y., Yang, K., Zhou, D., Qin, J., and Guo, X. (2010). Improving the Noah Land Surface Model in arid regions with an appropriate parameterization of the thermal roughness length. *Journal of Hydrometeorology*, 11(4):995–1006.

Choudhury, B. J. and DiGirolamo, N. E. (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 Hydrology*, 205(3):164–185.

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C., Swenson, S. C., and Zeng, X. (2015). Improving the representation of hydrologic processes in earth system models. *Water Resources Research*, 51(8):5929–5956.

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., and Luke, C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, 494:341.

Dai, Y., Dickinson, R. E., and Wang, Y.-P. (2004). A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, 17(12):2281–2299.

Deardorff, J. W. (1978). Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *Journal of Geophysical Research: Oceans*, 83(C4):1889–1903.

Dickinson, R. E. (1983). Land surface processes and climate—surface albedos and energy balance. *Advances in Geophysics*, 25:305–353. ID: 277243277243.

Dickinson, R. E., Kennedy, P., and Henderson-Sellers, A. (1993). *Biosphere-atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR community climate model*. National Center for Atmospheric Research, Climate and Global Dynamics Division, doi:10.5065/D67W6959, Boulder, CO.

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

Ehleringer, J. and Björkman, O. (1977). Quantum yields for CO_2 uptake in C3 and C4 plants: dependence on temperature, CO_2, and O_2 concentration. *Plant Physiology*, 59(1):86–90.

Ehleringer, J. and Pearcy, R. W. (1983). Variation in quantum yield for CO_2 uptake among C3 and C4 plants. *Plant Physiology*, 73(3):555–559.

Evans, J. (2013). Improving photosynthesis. *Plant Physiology*, 162(4):1780–1793.

Fan, Y., Meijide, A., Lawrence, D. M., Roupsard, O., Carlson, K. M., Chen, H.-Y., Röll, A., Niu, F., and Knohl, A. (2019). Reconciling canopy interception parameterization and rainfall forcing frequency in the Community Land Model for simulating evapotranspiration of rainforests and oil palm plantations in Indonesia. *Journal of Advances in Modeling Earth Systems*, 11(3):732–751.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta*, 149(1):78.

Fisher, J. B., Malhi, Y., Bonal, D., Rocha, H. R. D., Araujo, A. C. D., Gamo, M., Goulden, M. L., Hirano, T., Huete, A. R., Kondo, H., Kumagai, T., Loescher, H. W., Miller, S., Nobre, A. D., Nouvellon, Y., Oberbauer, S. F., Panuthai, S., Roupsard, O., Saleska, S., Tanaka, K., Tanaka, N., Tu, K. P., and Randow, C. V. (2009). The land–atmosphere water flux in the tropics. *Global Change Biology*, 15(11):2694–2714.

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J., Lichstein, J. W., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H., Powell, T. L., Serbin, S. P., Sato, H., Shuman, J. K., Smith, B., Trugman, A. T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., and Moorcroft, P. R. (2018). Vegetation demographics in earth system models: A review of progress and priorities. *Global Change Biology*, 24(1):35–54.

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., and Yu, Q. (2009). Simulation of within-canopy radiation exchange. *NJAS - Wageningen Journal of Life Sciences*, 57(1):5–15. ID: 278666.

Giambelluca, T. W., Martin, R. E., Asner, G. P., Huang, M., Mudd, R. G., Nullet, M. A., DeLay, J. K., and Foote, D. (2009). Evapotranspiration and energy balance of native wet montane cloud forest in Hawai'i. *Agricultural and Forest Meteorology*, 149(2):230–243.

Griend, A. A. and Owe, M. (1994). Bare soil surface resistance to evaporation by vapor diffusion under semiarid conditions. *Water Resources Research*, 30(2):181–188.

Gu, L., Shugart, H. H., Fuentes, J. D., Black, T. A., and Shewchuk, S. R. (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.

Gupta, S. and Larson, W. (1979). A model for predicting packing density of soils using particle-size distribution. *Soil Science Society of America Journal*, 43(4):758–764.

Haghighi, E. and Or, D. (2013). Evaporation from porous surfaces into turbulent airflows: Coupling eddy characteristics with pore scale vapor diffusion. *Water Resources Research*, 49(12):8432–8442.

Haghighi, E., Shahraeeni, E., Lehmann, P., and Or, D. (2013). Evaporation rates across a convective air boundary layer are dominated by diffusion. *Water Resources Research*, 49(3):1602–1610.

Harman, I. N. and Finnigan, J. J. (2007). A simple unified theory for flow in the canopy and roughness sublayer. *Boundary-Layer Meteorology*, 123(2):339–363. ID: Harman2007.

Harman, I. N. and Finnigan, J. J. (2008). Scalar concentration profiles in the canopy and roughness sublayer. *Boundary-Layer Meteorology*, 129(3):323–351. ID: Harman2008.

Hasler, N. and Avissar, R. (2007). What controls evapotranspiration in the amazon basin? *Journal of Hydrometeorology*, 8(3):380–395.

Heidkamp, M., Chlond, A., and Ament, F. (2018). Closing the energy balance using a canopy heat capacity and storage concept – a physically based approach for the land component JSBACHv3.11. 11(8):3465–3479.

Hillel, D. (2003). *Introduction to Environmental Soil Physics*. Academic Press, Burlington.

Huntingford, C., Zelazowski, P., Galbraith, D., Mercado, L. M., Sitch, S., Fisher, R., Lomas, M., Walker, A. P., Jones, C. D., Booth, B. B. B., Malhi, Y., Hemming, D., Kay, G., Good, P., Lewis, S. L., Phillips, O. L., Atkin, O. K., Lloyd, J., Gloor, E., Zaragoza-Castells, J., Meir, P., Betts, R., Harris, P. P., Nobre, C., Marengo, J., and Cox, P. M. (2013). Simulated resilience of tropical rainforests to CO_2-induced climate change. *Nature Geoscience*, 6(4):268–273.

Idso, S. (1974). The three stages of drying a field soil. *Soil Science Society of America Journal*, 38(5):831–837.

Jiménez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A., Fisher, J. B., Jung, M., Kanamitsu, M., Reichle, R. H., Reichstein, M., Rodell, M., Sheffield, J., Tu, K., and Wang, K. (2011). Global intercomparison of 12 land surface heat flux estimates. *Journal of Geophysical Research: Atmospheres*, 116.

Jung, M. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, 6(10):2001.

Katul, G., Manzoni, S., Palmroth, S., and Oren, R. (2010). A stomatal optimization theory to describe the effects of atmospheric co_2 on leaf photosynthesis and transpiration. *Annals of botany*, 105(3):431–442.

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

Kondo, J., Saigusa, N., and Sato, T. (1990). A parameterization of evaporation from bare soil surfaces. *Journal of Applied Meteorology*, 29(5):385–389.

Kozak, J. A., Ahuja, L. R., Ma, L., and Green, T. R. (2005). Scaling and estimation of evaporation and transpiration of water across soil textures. *Vadose Zone Journal*, 4(2):418–427.

Kume, T., Tanaka, N., Kuraji, K., Komatsu, H., Yoshifuji, N., Saitoh, T. M., Suzuki, M., and Kumagai, T. (2011). Ten-year evapotranspiration estimates in a Bornean tropical rainforest. *Agricultural and Forest Meteorology*, 151(9):1183–1192. ID: 271723.

Larsen, M. A. D., Refsgaard, J. C., Jensen, K. H., Butts, M. B., Stisen, S., and Mollerup, M. (2016). Calibration of a distributed hydrology and land surface model using energy flux measurements. *Agricultural and Forest Meteorology*, 217:74–88. ID: 271723.

Launiainen, S., Katul, G. G., Kolari, P., Vesala, T., and Hari, P. (2011). Empirical and optimal stomatal controls on leaf and ecosystem level CO_2 and H_2O exchange rates. *Agricultural and Forest Meteorology*, 151(12):1672–1689.

Launiainen, S., Katul, G. G., Lauren, A., and Kolari, P. (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.

Lawrence, D., Fisher, R., Koven, C., Oleson, K., Swenson, S., Vertenstein, M., Andre, B., Bonan, G., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Knox, R., Lawrence, P., Li, F., Li, H., Lombardozzi, D., Lu, Y., Perket, J., Riley, W., Sacks, W., Shi, M., Wieder, W., Xu, C., Ali, A., Badger, A., Bisht, G., Broxton, P., Brunke, M., Buzan, J., Clark, M., Craig, T., Dahlin, K., Drewniak, B., Emmons, L., Fisher, J., Flanner, M., Gentine, P., Lenaerts, J., Levis, S., Leung, L. R., Lipscomb, W., Pelletier, J., Ricciuto, D. M., Sanderson, B., Shuman, J., Slater, A., Subin, Z., Tang, J., Tawfik, A., Thomas, Q., Tilmes, S., Vitt, F., and Zeng, X., editors (2018). *Technical Description of version 5.0 of the Community Land Model (CLM)*. National Center for Atmospheric Research, Boulder, CO, ncar technical note edition. ID: 156.

Lawrence, D. and Vandecar, K. (2014). Effects of tropical deforestation on climate and agriculture. *Nature Climate Change*, 5:27.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence,

P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and 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).

Lawrence, P. J. and Chase, T. N. (2007). Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). *Journal of Geophysical Research: Biogeosciences*, 112(G1).

Lehmann, P. (2008). Characteristic lengths affecting evaporative drying of porous media. *Physical review.E, Statistical, nonlinear, and soft matter physics*, 77(5).

Lehmann, P., Merlin, O., Gentine, P., and Or, D. (2018). Soil texture effects on surface resistance to bare-soil evaporation. *Geophysical Research Letters*, 45(19):10,398–10,405.

Leuning, R. (2000). 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.

Li, X., Fan, X., and Brandani, S. (2014). Difference in pore contact angle and the contact angle measured on a flat surface and in an open space. *Chemical engineering science*, 117:137–145.

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

Mao, J., Ricciuto, D. M., Thornton, P. E., Warren, J. M., King, A. W., Shi, X., Iversen, C. M., and Norby, R. J. (2016). Evaluating the Community Land Model in a pine stand with shading manipulations and $13CO_2$ labeling. *Biogeosciences*, 13(3):641–657.

Massman, W. J. and Weil, J. C. (1999). 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.

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., Crous, K. Y., Angelis, P., Freeman, M., and Wingate, L. (2011). Reconciling the opti-

mal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, 17(6):2134–2144.

Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y., and Moorcroft, P. R. (2009). Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem demography model version 2. *Journal of Geophysical Research: Biogeosciences*, 114.

Mein, R. G. and Larson, C. L. (1973). Modeling infiltration during a steady rain. *Water Resources Research*, 9(2):384–394.

Merlin, O., Stefan, V. G., Amazirh, A., Chanzy, A., Ceschia, E., Er-Raki, S., Gentine, P., Tallec, T., Ezzahar, J., Bircher, S., Beringer, J., and Khabba, S. (2016). Modeling soil evaporation efficiency in a range of soil and atmospheric conditions using a meta-analysis approach. *Water Resources Research*, 52(5):3663–3684.

Michaletz, S. T., Weiser, M. D., McDowell, N. G., Zhou, J., Kaspari, M., Helliker, B. R., and Enquist, B. J. (2016). The energetic and carbon economic origins of leaf thermoregulation. *Nature Plants*, 2:16129.

Miller, G., Moore, G., Cahill, A., Aparecido, L., Andrews, R., and Song, J. (2018a). Texas A&M soltis center forest canopy data, 2014. `http://hdl.handle.net/1969.1/169521.2`.

Miller, G., Moore, G., Cahill, A., Aparecido, L., Andrews, R., and Song, J. (2018b). Texas A&M soltis center forest canopy data, 2015. `http://hdl.handle.net/1969.1/169522`. ID: 282.

Miller, G., Moore, G., Cahill, A., Aparecido, L., Andrews, R., and Song, J. (2018c). Texas A&M soltis center forest canopy data, 2016. `http://hdl.handle.net/1969.1/169523`. ID: 282.

Miller, G., Moore, G., Cahill, A., Aparecido, L., Andrews, R., and Song, J. (2018d). Texas A&M soltis center forest canopy data, 2017. `http://hdl.handle.net/1969.1/169524`. ID: 282.

Monteith, J. and Unsworth, M. (2013). *Principles of Environmental Physics : Plants, Animals, and the Atmosphere*. Academic Press.

Moore, G. W., Orozco, G., Aparecido, L. M. T., and Miller, G. R. (2018). Upscaling transpiration in diverse forests: Insights from a tropical premontane site. *Ecohydrology*, 11(3):e1920.

Mosthaf, K. (2014). Modeling and analysis of evaporation processes from porous media on the rev scale. *Water Resources Research*, 50(2):1059–1079.

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, 12(3):513–522.

Muñoz-Villers, L. E., Holwerda, F., Gómez-Cárdenas, M., Equihua, M., Asbjornsen, H., Bruijnzeel, L. A., Marín-Castro, B. E., and Tobón, C. (2012). Water balances of old-growth and regenerating montane cloud forests in central Veracruz, Mexico. *Journal of Hydrology*, 462-463:53–66.

Neriah, A. B. (2014). Impact of ambient conditions on evaporation from porous media. *Water Resources Research*, 50(8):6696–6712.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and 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).

Ogée, J., Brunet, Y., Loustau, D., Berbigier, P., and Delzon, S. (2003). 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.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Kluzek, E., Lamarque, J. F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D. M., Sacks, W., Sun, Y., Tang, J., and Yang, Z. L., editors (2013). *Technical*

Description of Version 4.5 of the Community Land Model (CLM). National Center for Atmospheric Research, Boulder, CO, ncar technical note ncar/tn-503 str edition.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J. F., Mahowald, N., Niu, G. Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z. L., Zeng, X., and Zeng, X., editors (2010). *Technical description of version 4.0 of the Community Land Model (CLM)*. National Center for Atmospheric Research, Boulder, CO, ncar technical note ncar/tn-478+str edition.

Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T. (2008). Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research: Biogeosciences*, 113(G1).

Papale, D. (2006). Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences*, 3(4):571.

Philip, J. R. (1957). Evaporation, and moisture and heat fields in the soil. *Journal of Meteorology*, 14(4):354–366.

Pitman, A. J., de Noblet-Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., Claussen, M., Delire, C., Ganzeveld, L., Gayler, V., den Hurk van, Lawrence, P. J., der Molen van, Müller, C., Reick, C. H., Seneviratne, S. I., Strengers, B. J., and Voldoire, A. (2009). Uncertainties in climate responses to past land cover change: First results from the lucid intercomparison study. *Geophysical Research Letters*, 36(14).

Qian, T., Dai, A., Trenberth, K. E., and Oleson, K. W. (2006). Simulation of global land surface conditions from 1948 to 2004. part I: forcing data and evaluations. *Journal of Hydrometeorology*, 7(5):953–975.

Raj, R., Hamm, N. A. S., van der Tol, C., and Stein, A. (2015). Uncertainty analysis of gross

primary production partitioned from net ecosystem exchange measurements. *Biogeosciences discussions*, 12(16):13967–14002.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, 11(9):1424–1439.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, 1(5):318–333.

Rivera, M. A., Keebler, A., Smith, A. P., Miller, G. R., Aitkenhead-Peterson, J. A., and Moore, G. W. (2018). The effect of tropical land use on soil carbon dynamics: Does reforestation mitigate greenhouse gas emissions? In *AGU Fall Meeting Abstracts*.

Ryder, J., Polcher, J., Peylin, P., Ottlé, C., Chen, Y., van Gorsel, E., Haverd, V., McGrath, M. J., Naudts, K., Otto, J., Valade, A., and 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.

Sadeghi, M., Shokri, N., and Jones, S. B. (2012). A novel analytical solution to steady-state evaporation from porous media. *Water Resources Research*, 48(9).

Sadeghi, M., Tuller, M., Gohardoust, S., and Jones (2014). Column-scale unsaturated hydraulic conductivity estimates in coarse-textured homogeneous and layered soils derived under steady-state evaporation from a water table. *Journal of hydrology*, 519:1238–1248.

Sakaguchi, K. and Zeng, X. (2009). Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3.5). *Journal of Geophysical Research: Atmospheres*, 114(D1).

Schlesinger, W. H. and Jasechko, S. (2014). Transpiration in the global water cycle. *Agricultural and Forest Meteorology*, 189:115–117. ID: 271723.

Schlosser, C. A. and Gao, X. (2010). Assessing evapotranspiration estimates from the Second Global Soil Wetness Project (GSWP-2) simulations. *Journal of Hydrometeorology*, 11(4):880–897.

Schlünder, E. U. (1988). On the mechanism of the constant drying rate period and its relevance to diffusion controlled catalytic gas phase reactions. *Chemical Engineering Science*, 43(10):2685–2688.

Sellers, P., Berry, J., Collatz, G., Field, C., and 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.

Shahraeeni, E., Lehmann, P., and Or, D. (2012). Coupling of evaporative fluxes from drying porous surfaces with air boundary layer: Characteristics of evaporation from discrete pores. *Water Resources Research*, 48(9).

Sheil, D. (2018). Forests, atmospheric water and an uncertain future: the new biology of the global water cycle. *Forest Ecosystems*, 5(1):19. ID: Sheil2018.

Sherwood, T. K. (1930). The drying of Solids-III[1] mechanism of the drying of pulp and paper. *Industrial & Engineering Chemistry*, 22(2):132–136.

Shokri, N., Lehmann, P., Vontobel, P., and Or, D. (2008). Drying front and water content dynamics during evaporation from sand delineated by neutron radiography. *Water Resources Research*, 44(6).

Shokri, N. and Salvucci, G. D. (2011). Evaporation from porous media in the presence of a water table. *Vadose Zone Journal*, 10(4):1309–1318.

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

Singh, R. S., Reager, J. T., Miller, N. L., and Famiglietti, J. S. (2015). Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations over the southwestern U.S. *Water Resources Research*, 51(4):2648–2667.

Skillman, J. B. (2007). Quantum yield variation across the three pathways of photosynthesis: not yet out of the dark. *Journal of experimental botany*, 59(7):1647–1661.

Song, J., Miller, G. R., Cahill, A. T., Aparecido, L. M. T., and Moore, G. W. (2019). Modeling land surface processes over a mountainous rainforest in costa rica using clm4.5 and clm5. In preparation for submission to Geoscientific Model Development.

Suzuki, M. and Maeda, S. (1968). On mechanism of drying of granular beds. *Journal of Chemical Engineering of Japan*, 1(1):26–31.

Swenson, S. C. and 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.

Teale, N. G., Mahan, H., Bleakney, S., Berger, A., Shibley, N., Frauenfeld, O. W., Quiring, S. M., Rapp, A. D., Roark, E. B., and 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.

Teng, J., Yasufuku, N., Liu, Q., and Liu, S. (2014). Experimental evaluation and parameterization of evaporation from soil surface. *Natural Hazards*, 73(3):1405–1418.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., and Mahowald, N. M. (2007). Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. *Global Biogeochemical Cycles*, 21(4).

Tschapek, M., Scoppa, C. O., and Wasowski, C. (1978). The surface tension of soil water. *Journal of Soil Science*, 29(1):17–21.

van Genuchten, M. T. (1980). Closed-from equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5):892–898.

Vose, J. M., Clinton, B. D., Sullivan, N. H., and Bolstad, P. V. (1995). Vertical leaf area distribution, light transmittance, and application of the Beer-Lambert law in four mature hardwood stands in the southern Appalachians. *Canadian Journal of Forest Research*, 25(6):1036–1043.

Wang, A., Barlage, M., Zeng, X., and 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.

Wang, A., Zeng, X., and Guo, D. (2016). Estimates of global surface hydrology and heat fluxes from the Community Land Model (CLM4.5) with four atmospheric forcing datasets. *Journal of Hydrometeorology*, 17(9):2493–2510.

Wang, A., Zeng, X., Shen, S. S. P., Zeng, Q.-C., and Dickinson, R. E. (2006). Time scales of land surface hydrology. *Journal of Hydrometeorology*, 7(5):868–879.

Wohl, E., Barros, A., Brunsell, N., Chappell, N. A., Coe, M., Giambelluca, T., Goldsmith, S., Harmon, R., Hendrickx, J. M. H., Juvik, J., McDonnell, J., and Ogden, F. (2012). The hydrology of the humid tropics. *Nature Climate Change*, 2:655.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring earth's terrestrial water. *Water Resources Research*, 47(5).

Yang, K., Koike, T., Fujii, H., Tamagawa, K., and Hirose, N. (2002). Improvement of surface flux parametrizations with a turbulence-related length. *Quarterly Journal of the Royal Meteorological Society*, 128(584):2073–2087.

Yiotis, A. G., Tsimpanogiannis, I. N., Stubos, A. K., and Yortsos, Y. C. (2006). Pore-network study of the characteristic periods in the drying of porous materials. *Journal of colloid and interface science*, 297(2):738–748.

Zaehle, S. and Friend, A. D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface

model: 1. model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, 24(1).

Zeng, X., Wang, Z., and Wang, A. (2012). Surface skin temperature and the interplay between sensible and ground heat fluxes over arid regions. *Journal of Hydrometeorology*, 13(4):1359–1370.

Zeng, X., Zhao, M., and Dickinson, R. E. (1998). Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *Journal of Climate*, 11(10):2628–2644.

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

Zhao, W. and Qualls, R. J. (2005). A multiple-layer canopy scattering model to simulate shortwave radiation distribution within a homogeneous plant canopy. *Water Resources Research*, 41(8).

Zheng, W., Wei, H., Wang, Z., Zeng, X., Meng, J., Ek, M., Mitchell, K., and Derber, J. (2012). Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation. *Journal of Geophysical Research: Atmospheres*, 117(D6).

APPENDIX A

HYPERGEOMETRIC FUNCTION

The Gaussian hypergeometric function 2F1 can be defined as

$$2F1(a_f, b_f, c_f, z) = \sum_{n=0}^{\infty} \frac{(a_f)_n (b_f)_n}{n!(c_f)_n} z^n \tag{A.1}$$

where any $(x)_n$ is

$$(x)_n = \begin{cases} 1 & if \;\; n = 0 \\ x(x+1)(x+2)\cdots(x+n-1) & if \;\; n > 0 \end{cases} \tag{A.2}$$

When we consider a part of Eq. (2.14) as

$$\sum_{n=0}^{\infty} \frac{h_{surf}}{Pn+1} \left[-\frac{e_0}{c \cdot K_s} \left(\frac{h_{surf}}{h_b} \right)^P \right]^n \tag{A.3}$$

To convert the Maclaurin series expansion of Eq. (A.3) into a hypergeometric function form, we first define z as equal to $-\frac{e_0}{c \cdot K_s} \left(\frac{h_{surf}}{h_b} \right)^P$. To set $1/(Pn+1)$ in a form of $\frac{(a_f)_n (b_f)_n}{n!(c_f)_n}$, and simply define $a = 1$. We also note that for $c = b + 1$ we have

$$\frac{(b_f)_n}{(c_f)_n} = \frac{b_f (b_f+1)(b_f+2)\cdots(b_f+n-1)}{(b_f+1)(b_f+2)(b_f+3)\cdots(b_f+n)} = \frac{b_f}{b_f+n} \tag{A.4}$$

Therefore, with b_f equal to $1/p$, c_f becomes $1/p + 1$, which provides the final relationship as

$$\sum_{n=0}^{\infty} \frac{h_{surf}}{Pn+1} \left[-\frac{e_0}{c \cdot K_s} \left(\frac{h_{surf}}{h_b} \right)^P \right]^n = h_{surf} \cdot 2F1 \left(1, \frac{1}{P}, 1 + \frac{1}{P}, -\frac{e_0}{c \cdot K_s} \left(\frac{h_{surf}}{h_b} \right)^P \right) \tag{A.5}$$

APPENDIX B

UPDATED CALCULATIONS FOR MULTI-LAYERED CLM

B.1 Additional Plots

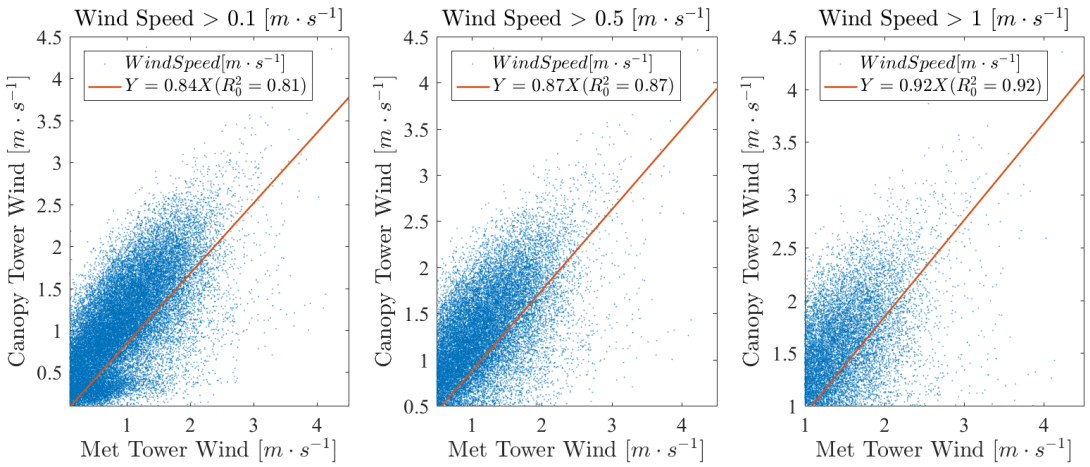

Figure B.1: Comparison plots for wind speed data between Canopy Tower and Met Tower. Slope and its correlation get higher when wind speed is higher. CLM uses wind speed bigger than $1\mathrm{m}\cdot\mathrm{s}^{-1}$. The fitting is only for a slope without intercept so the R-squared R_0^2 is normally higher than usual case.

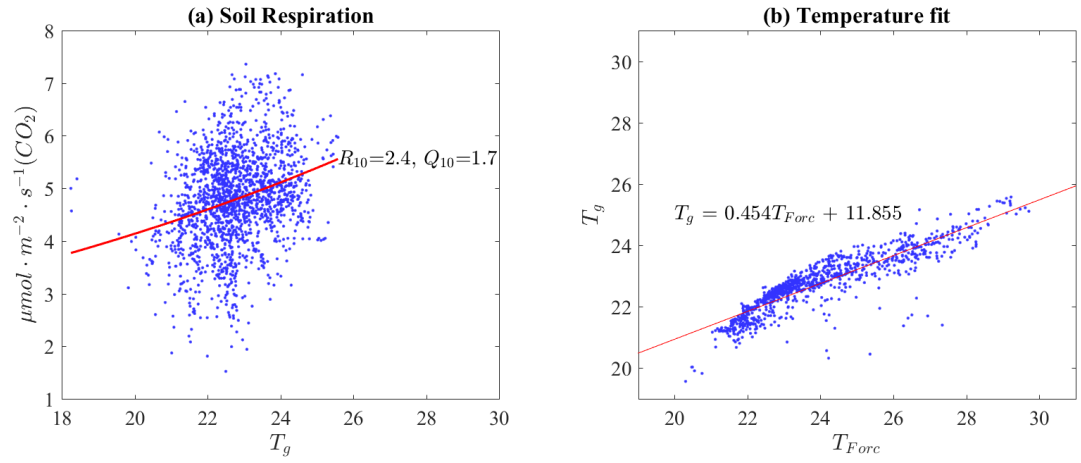

Figure B.2: (a) Soil respiration data from Rivera et al. (2018) and fitted parameters. (b) Temperature fitting between forcing temperature and ground temperature.

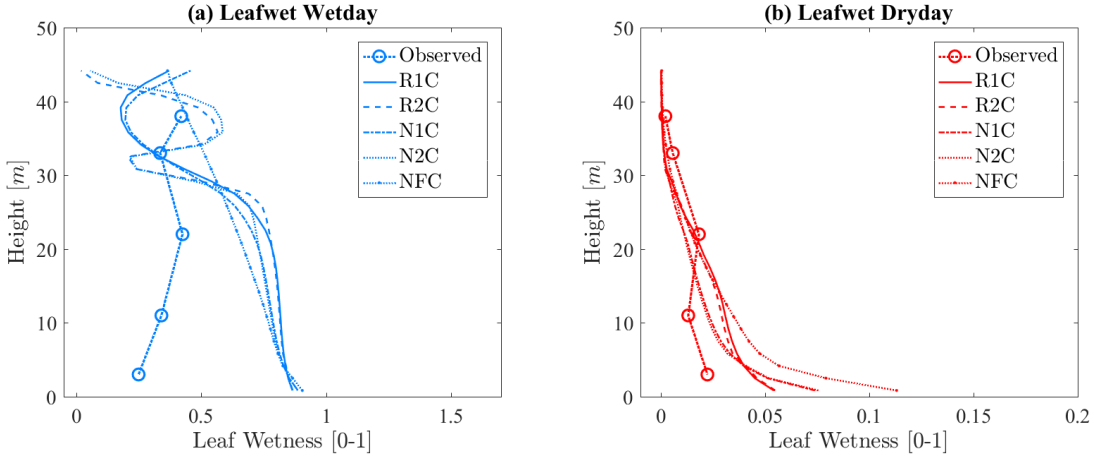

Figure B.3: Leaf wetness profile.

B.2 Additional Description of Wind Profile Models and CO_2 Profile Models

B.2.1 Wind Speed Profile Model: First-Order Closure Model

First-order closure model is solved using a numerical method. The model Eq. (5.4) follows other literature (Launiainen et al., 2011; Katul et al., 2004; Drewry et al., 2010), which has

$$K_m \frac{\partial^2 u}{\partial z^2} + \frac{\partial K_m}{\partial z}\frac{\partial u}{\partial z} - C_d a(z) u^2 = 0 \tag{B.1}$$

where K_m is the eddy diffusivity for momentum, u is wind speed, z is height, C_d is drag coefficient (0.25), and $a(z)$ is LAD (Launiainen et al., 2011). The eddy diffusivity for momentum is

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

$$\frac{\partial^2 u}{\partial z^2} = \frac{u_{i-1} - 2u_i + u_{i+1}}{dz^2} \tag{B.2}$$

And the first-order derivative is

$$\frac{\partial u}{\partial z} = \frac{u_{i-1} - u_{i+1}}{2dz} \tag{B.3}$$

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

$$\left|\frac{\partial u}{\partial z}\right|_i = \left|\frac{u_{i-1} - u_i}{dz}\right| \tag{B.4}$$

The eddy diffusivity for momentum like Eq. (5.5) is

$$K_{m,i} = l_{m,i}^2 \left|\frac{\partial u}{\partial z}\right|_i \tag{B.5}$$

where l is mixing length. Following Launiainen et al. (2011) and Katul et al. (2004), it is like Eq. (5.6) as

$$l_m = \begin{cases} k_v z, & z < \alpha' H_{max}/k_v \\ \alpha' H_{max}, & \alpha' H_{max}/k_v \le z < H_{max} \\ k_v(z-d), & H_{max} \le z \end{cases} \tag{B.6}$$

where k_v is Von Karman constant, α' is $k_v(1\breve{}d/H_{max})$, 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 $0.667 \cdot H_{max}$ but it varies in CLMML (Bonan et al., 2018). Then, the derivative of the eddy diffusivity is

$$\frac{\partial K_{m,i}}{\partial z} = \frac{K_{m,i} - K_{m,i-1}}{dz} \tag{B.7}$$

If wind value and its vertical gradient are known as an initial condition, it becomes the simple ordinary differential equations (ODEs). The case of the wind profile above the EC location is the such case, so it can be estimated using various well-known methods such as Euler method, Midpoint method, etc. (we used Midpoint method for this study). On the other hand, for the below the EC location, we only know wind values at the bottom and at the EC location. In that case, another method such as Tridiagonal matrix solution is necessary to solve the equation as below. After solving the wind profile below the EC location, the required initial values of the ODE solution are obtained for the higher locations.

For the solution of the below the EC location, applying all the numerical forms in the first-order closure model, the final relationship becomes

$$\begin{aligned} K_{m,i}u_{i-1} - (K_{m,i} + K_{m,i+1})u_i + K_{m,i+1}u_{i+1} + \\ 0.5 \cdot (K_{m,i} - K_{m,i-1}) \cdot (u_{i-1} - u_{i+1}) - \\ C_d a(z) u_i u_i \cdot dz^2 = 0 \end{aligned} \tag{B.8}$$

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

$$\begin{aligned} [K_{m,i} + 0.5 \cdot (K_{m,i} - K_{m,i+1})] \cdot u_{i-1} \\ -[(K_{m,i} + K_{m,i+1}) + C_d a(z) u_i \cdot dz^2] \cdot u_i \\ +[K_{m,i+1} - 0.5 \cdot (K_{m,i} - K_{m,i+1})] \cdot u_{i+1} = 0 \end{aligned} \tag{B.9}$$

Usage of the conductivity at the ground g_0 in CLM5 was also tried to apply it as a lower boundary condition for Km but it constantly produces negative wind speed at the near ground. Therefore, $g_{a,0}$ which has around $g_{a,0} \approx 0.0768$ in CLM5 was not applied for the first-order closure model.

B.2.2 Wind Speed Profile Model: Roughness Sublayer (RSL) Model

The tracking the wind profile using RSL model is straightforward because it is an analytical solution that the function consists of other than the wind speed ($u(z) = f(z)$). The RSL model

equation for $z > H_{max}$ is

$$u(z) = \frac{u^*}{k_v}\left[ln\left(\frac{z-d}{H_{max}-d}\right) - \psi_m\left(\frac{z-d}{L_{MO}}\right) + \psi_m\left(\frac{H_{max}-d}{L_{MO}}\right) + \hat{\psi_m}\left(\frac{z-d}{L_{MO}}, \frac{z-d}{l_M/\beta}\right) - \hat{\psi_m}\left(\frac{H_{max}-d}{L_{MO}}, \frac{H_{max}-d}{l_M/\beta}\right) + \frac{k_v}{\beta}\right] \tag{B.10}$$

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, $\hat{\psi}_m$ 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 and Finnigan, 2008, 2007). Also, The RSL model equation for $z \leq H_{max}$ is

$$u(z) = u(h_{max}) \exp\left[\frac{z - H_{max}}{l_M/\beta}\right] \tag{B.11}$$

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

$$u(z) = u_{ref}(z_{ref}) \cdot f(z_{ref})/f(z) \tag{B.12}$$

B.2.3 CO_2 Profile Model

The method to estimate CO_2 profile is similar to other profile estimations like as RH or air temperature, described by Bonan et al. (2018). The Eq. (5.8) is

$$\rho_m \frac{\partial C}{\partial t} - \frac{\partial}{\partial z}\left(\rho_m K_c(z) \frac{\partial C}{\partial z}\right) = \left[f_{c,sun}(z) f_{sun}(z) + f_{c,sha}(z) f_{sha}(z)\right] a(z) \tag{B.13}$$

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, 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

$$\rho_m \frac{\Delta z}{\Delta t}\left(C_i^{t+1} - C_i^t\right) - g_{a,i-1}\left(C_{i-1}^{t+1} - C_i^{t+1}\right) + g_{a,i}\left(C_{i-1}^{t+1} - C_{i+1}^{t+1}\right) = f_{c,i}^{t+1} L_i(z) \tag{B.14}$$

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

$$\begin{aligned} & -g_{a,i-1}C_{i-1}^{t+1} \\ & \left(\rho_m \frac{\Delta z}{\Delta t} + g_{a,i-1} + g_{a,i}\right) C_i^{t+1} \\ & -g_{a,i}C_{i+1}^{t+1} \\ & = f_{c,i}^{t+1} L_i(z) + \rho_m \frac{\Delta z}{\Delta t} C_i^t \end{aligned} \tag{B.15}$$

B.2.4 LAD Profile Model

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

$$f_{LAD,1}(z, H) = \frac{L}{H} \cdot f_{Beta}(z/H, p, q) + \frac{S}{H} \tag{B.16}$$

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 LAD model can be written using mixed-distribution as

$$\begin{aligned} f_{LAD,2}(z, H_d, r, \Delta H) = r \cdot f_{LAD,2}(z + (r-1)\cdot\Delta H, H_d) + \\ (1-r)\cdot LAD_1(z + r\cdot\Delta H, H_d) \end{aligned} \tag{B.17}$$

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 5.1]. 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

$$f_{LAD,2,2}(z, H_d, r_d, \Delta H_d, H_m, r_m, \Delta H_m, v) = v \cdot f_{LAD,2}(z, H_d, r_d, \Delta H_d) + (1-v) \cdot f_{LAD,2}(z, H_m, r_m, \Delta H_m) \tag{B.18}$$

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 data (Song et al., 2019), using the least-squares method. Here, for proper fitting, all LADs are converted in a cumulative form because the main purpose 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 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 B.1].

Table B.1: 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

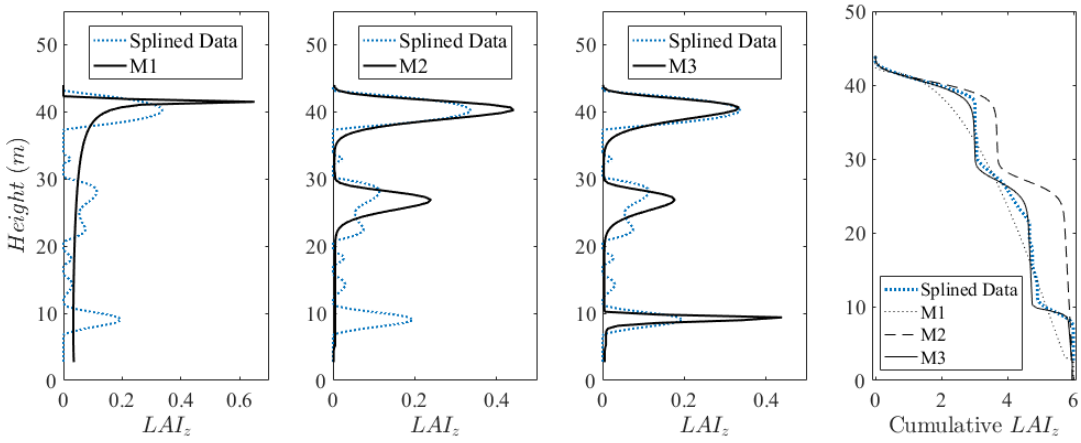

Figure B.4: LAD profile test.

APPENDIX C

SUPPLEMENTAL SOURCES CONSULTED

- Microsoft Excel File
  - Containing : This file is the simulation result from all listed setups in [Table 5.1]. The file provides the total fluxes (e.g., net radiation, carbon dioxide, latent heat flux, and heat flux) and the up-scaled variables (e.g., leaf wetness and vegetation temperature) of observations and simulations in Section 5. Also, their difference, observations minus simulations, are calculated and plotted. The data were categorized into day time, wet day time, dry day time, night time.
  - File name : "Supplement.xlsx" (82KB)