Estimation of global land surface evapotranspiration with the consideration of vegetation structural and physiological status from remote sensing

by

Xiangzhong Luo

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Geography and Planning
University of Toronto

© Copyright by Xiangzhong Luo (2018)
Global land surface evapotranspiration (ET) is a critical constituent of the global energy and water budgets. The average of perennial terrestrial ET accounts for around 60% of precipitation and consumes about 50% of the absorbed solar radiation on the land surface. At least 50% of terrestrial ET is comprised of transpiration – the water efflux through leaf stomata – the variations of which are subject to vegetation status. The structural status of vegetation determines the radiation distribution inside a canopy and the available energy for transpiration and photosynthesis, while physiological status describes the activity of pigments and enzymes associated with the biochemical processes that control photosynthesis and stomata. Currently, the status of vegetation is often underrepresented in terrestrial biosphere models (TBMs) – the principle tools used to simulate terrestrial ET, due to the difficulty of direct measurements of vegetation traits over large scales. Chapter 2 uses ground measured leaf area index (LAI) and clumping index (CI) over nine flux towers to form several upscaling schemes for a TBM and suggests that a two-leaf scheme (TL) is the best upscaling scheme for modelling ET, based on a comprehensive comparison. TL captures the instantaneous features of the radiation distribution and avoids uncertainties in calculating canopy-scale biochemical parameters. Chapter 3 incorporates the ground measured leaf chlorophyll content (LCC)
from a forest site into a TBM and establishes a framework to use LCC in TL. Chapter 3 reveals that LCC can work as a robust proxy for the leaf photosynthetic capacity in TBM, with the estimates of gross primary productivity (GPP) and ET in being considerably improved in spring and autumn. After the site-level study, Chapter 4 uses satellite-derived LAI, CI and LCC to simulate ET and GPP at 124 sites across nine plant functional types distributed globally. Satellite-derived LCC can reduce the biases of ET and GPP estimates at around 60% of the 554 site-years, with the greatest reduction for plant functional types with strong seasonal cycles. This study demonstrates the critical role of vegetation in simulating global ET and GPP, and develops the optimal way to incorporate vegetation structural and physiological parameters in ET modelling.
Acknowledgement

Even though my name appears alone on the front page of this thesis, it would not be possible for me to finish it without the help from a large number of people, to whom I want to express my earnest gratitude here.

First, I would like to thank my supervisor, Prof. Jing M. Chen. I still remember the excitement when Jing first replied my email saying he would consider accepting me as a PhD student. In our first meeting, Jing warned me that PhD is a painstaking but fun process, and I should stay ready for the “pain” part. I gradually understand its meaning as I am approaching the finish line, and luckily, I still feel the fun to continue research. Just like that case, Jing always gives me the honest and unbiased suggestions under every circumstance. Jing shows me the passion, mindset and ethic that a top scientist should embody and inspires me to seek scientist as a career. Although Jing is obligated to many academic duties, he can always manage to provide me strong supports and critical guidance at every stage of my PhD.

I am also very grateful to my thesis committee: Prof. Jane Liu, Prof. Danny Harvey, Prof. Paul Kushner, Prof. Sarah Finkelstein and Prof. Dennis Baldocchi. One cannot find a more encouraging professor than Jane. Jane persuaded me to focus on ET in the first place when I got lost in the wide spectrum and depth of Jing’s publications in my first year here. Danny helped me with critical thinking and academic writing through meticulously reviewing a lot of documents written by me. Paul is kind but not shy of giving me hard questions. I guess every time I think of Paul, his question of “what is your model?” will resonate in my mind and alert me to think from a physicist’s perspective. Sarah kindly provided us a spectrophotometer for many years to measure leaf chlorophyll, which makes the Chapters 3 and 4 of this thesis possible. She gave me a lot of suggestions on lab safety and management when I organized the group lab work in year 3. It is also my great pleasure to have Prof. Baldocchi as an external examiner. His papers have helped me through my project in the past 4 years and his comments really show me the big-picture of global carbon and water studies. As a well-respected and
senior scientist in our community, Prof. Baldocchi always shows curiosity to new ideas and patience to young scientists.

Additionally, I would like to extend a special thanks to Dr. Holly Croft, one of our group’s valuable postdocs, who acts as my day-to-day research mentor. Holly patiently walks me through field work, lab work and paper writing, and encourages me to take on challenging topics. I would really miss the time with her in the field trips and our broad discussions on numerous academic and non-academic stuffs. I also appreciate the help from Dr. Alemu Gonsamo, Dr. Liming He and our out-going technician Gang Mo, for their patience to discuss and to work with me. In addition, a massive thanks to my adorable lab mates, including but not limited to, Dr. Ting Zheng, Dr. Rong Wang, Nadine Nesbit, Ze Wang, Joyce Arabian, Jenny Jung, Parastoo Hossein and Cheryl Rogers. They have brought enormous fun to our group and let me feel I am not alone in this process. I am also thankful to Dr. Ralf Staebler and Dr. Paul Bartlett from Environment Canada, Prof. Altaf Arain from McMaster University, Prof. Marney Isaac from University of Toronto Scarborough and Newton Tran from University of Guelph for their support of my field work. I also want to thank the data producers and distributors of FLUXNET and MERIS.

At last, I am grateful to my friends in Toronto and Calgary for providing numerous happy hours: playing cards buddies, ex-roommates, fellow graduate students, physics geeks, my girlfriend’s girlfriend’s boyfriend. Their company made this journey more memorable. I want to thank my folks, for whatever I do, they are my hard-core supporters. My father, who is a retired pharmacist, even tried to understand the terrestrial carbon and water cycle. I want to thank Dr. Bing Yang, my newly married wife and the girl I like the most in the world. You are smart, kind and beautiful and always believe in me more than myself.

University of Toronto, the Natural Sciences and Engineering Research Council of Canada (NSERC) and Jing generously provide the financial support for me to study and live in Canada. This thesis is a comma in my journey at the age of 28. I am looking forward to the new challenges in the next chapter.
# Table of Contents

Acknowledgement .................................................................................................................. iv
Table of Contents .................................................................................................................... vi
List of Tables .......................................................................................................................... x
List of Figures .......................................................................................................................... xi
Chapter 1 ................................................................................................................................. 1
  1.1 Characteristics of terrestrial evapotranspiration .............................................................. 1
  1.1.1 Global land surface ET .............................................................................................. 1
  1.1.2 Trends and variability of ET ....................................................................................... 5
  1.1.3 Transpiration and water use efficiency ........................................................................ 6
  1.2 Measurement and modelling of ET .................................................................................. 10
  1.2.1 Eddy covariance technique ....................................................................................... 10
  1.2.2 Modelling of ET ........................................................................................................... 13
  1.2.3 Modelling stomatal controls ...................................................................................... 16
  1.3 Remote sensing of vegetation traits ................................................................................ 18
  1.3.1 Vegetation structural and physiological traits .............................................................. 18
  1.3.2 Remote sensing of physiological traits ........................................................................ 20
  1.4 Boreal Ecosystem Productivity Simulator ....................................................................... 24
  1.4.1 Definition of Terrestrial Biosphere models (TBM)s .................................................. 24
  1.4.2 Development of BEPS .............................................................................................. 26
  1.5 Roadmap of the thesis .................................................................................................... 27
  1.6 References ...................................................................................................................... 28
Chapter 2 ................................................................................................................................. 44
Comparison of big-leaf, two-big-leaf and two-leaf upscaling schemes for evapotranspiration estimation using coupled carbon-water modelling ................................................. 44
  2.1 Abstract .......................................................................................................................... 44
  2.2 Introduction ..................................................................................................................... 45
2.3 Data and method ................................................................................................. 49
   2.3.1 Description of the model ........................................................................... 49
   2.3.2 Modelling schemes .................................................................................. 51
   2.3.3 Validation sites and input data .................................................................. 53
2.4 Results ................................................................................................................ 57
   2.4.1 Annual ET and GPP estimates from each scheme ..................................... 57
   2.4.2 Biases of ET and GPP estimates ............................................................... 59
   2.4.3 Difference between the radiation regimes in BL and TBL(TL) ............... 60
   2.4.4 Difference between TBL and TL ............................................................... 62
2.5 Discussion .......................................................................................................... 66
2.6 Conclusion .......................................................................................................... 68
2.7 Appendix ............................................................................................................ 69
   2.7.1 Nitrogen-weighted Vmax25 and Jmax25 for sunlit and shaded leaves ...... 69
   2.7.2 Leaf energy budget .................................................................................... 71
   2.7.3 Quantification of the soil water stress factor ........................................... 75
   2.7.4 Parameterization for TBL ......................................................................... 75
   2.7.5 Correlations between the simulated hourly ET (GPP) and measured hourly ET (GPP) under all schemes. .......................................................... 77
2.8 References ......................................................................................................... 79

Chapter 3 .................................................................................................................. 86
Incorporating leaf chlorophyll content into a two-leaf terrestrial biosphere model for estimating carbon and water fluxes at a forest site .................................................. 86
   3.1 Abstract .......................................................................................................... 86
   3.2 Introduction .................................................................................................... 87
   3.3 Materials and methods .................................................................................. 90
      3.3.1 Site profile ................................................................................................ 90
      3.3.2 Field measurements ............................................................................... 91
      3.3.3 Flux and meteorological measurements ................................................ 92
      3.3.4 Terrestrial biophysical model .................................................................. 94
      3.3.5 Modelling treatments .............................................................................. 98
Chapter 3

3.4. Results ......................................................................................................................... 101

3.4.1 Leaf traits and canopy fluxes .................................................................................. 101

3.4.2 Validation of the GPP and ET estimates from different Vcmax25 constraints ............. 103

3.4.3 Seasonality of the improvements in estimated GPP and ET .................. 104

3.4.4 Assessing the relative contribution of Chl_leaf to modelled GPP and ET ........ 106

3.4.5 The impact of light environment within the canopy on leaf physiology 107

3.5. Discussion .................................................................................................................. 111

3.5.1 Improved seasonal representation for GPP and ET simulations ................. 111

3.5.2 Physiological differences between sunlit and shaded leaves ..................... 112

3.6. Conclusions .............................................................................................................. 114

3.7 Appendix ..................................................................................................................... 116

3.7.1 Net radiation on sunlit and shaded leaves .................................................. 116

3.7.2 Quantification of the soil water stress factor ............................................. 118

3.8 References .................................................................................................................. 119

Chapter 4

Chlorophyll dynamics explains spatial and temporal variabilities in terrestrial carbon uptake and evapotranspiration ................................................................. 127

4.1 Abstract ...................................................................................................................... 127

4.2 Introduction ............................................................................................................... 128

4.3 Materials and method ............................................................................................. 130

4.3.1 Flux and meteorological dataset ...................................................................... 130

4.3.2 Remote sensing data ......................................................................................... 132

4.3.3 Terrestrial biosphere model .............................................................................. 134

4.3.4 Derivation of LCC ............................................................................................. 136

4.3.5 Derivation of Vcmax25 from LCC .................................................................... 136

4.4 Results ....................................................................................................................... 137

4.4.1 Seasonality of LCC and Vcmax25 according to PFT ................................ 137

4.4.2 Examining the advantage of using LCC to model GPP and ET .............. 140

4.5 Discussion ................................................................................................................. 144
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5.1 Assessment of the Vcmax25-LCC relationships</td>
<td>144</td>
</tr>
<tr>
<td>4.5.2 Uncertainty of LCC for evergreen PFTs</td>
<td>146</td>
</tr>
<tr>
<td>4.5.3 Relative sensitivities of GPP and ET fluxes to biophysical and environmental drivers</td>
<td>148</td>
</tr>
<tr>
<td>4.6 Conclusion</td>
<td>149</td>
</tr>
<tr>
<td>4.7 Appendix</td>
<td>150</td>
</tr>
<tr>
<td>4.7.1 Validation of BEPS with and without LCC</td>
<td>150</td>
</tr>
<tr>
<td>4.7.3 Energy closure problem</td>
<td>157</td>
</tr>
<tr>
<td>4.8 References</td>
<td>159</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>167</td>
</tr>
<tr>
<td>Summary</td>
<td>167</td>
</tr>
<tr>
<td>5.1. Lessons learned from this research</td>
<td>167</td>
</tr>
<tr>
<td>5.2. Topics for further research</td>
<td>168</td>
</tr>
</tbody>
</table>
List of Tables

Table 1-1. Global land surface evapotranspiration (ET)………………………………………2

Table 1-2. Component of global terrestrial ET…………………………………………………7

Table 2-1. Features of the selected Canada Carbon Program (CCP) flux-tower sites………………………………………………………………………………………………………55

Table 2-A1. Correlations between the simulated hourly ET (GPP) and measured hourly ET (GPP) under all schemes………………………………………………………………………77

Table 3-1. Slopes in the linear Chl_{leaf}-V_{\text{max}}^{25} relationships reported in the literatures…………………………………………………………………………………………………………………………113

Table 4-1. Linear regression equations used to derive V_{\text{max}}^{25} from LCC………………137

Table 4-2. The mean and standard deviations of site-average V_{\text{max}}^{25} for each PFTs………………………………………………………………………………………………………………139

Table 4-3. V_{\text{max}}^{25}-LCC relationships (V_{\text{max}}^{25} = a*LCC + b) derived using different methods………………………………………………………………………………………………………………145

Table 4-A1. Linear correlations between the daily estimated fluxes and measured fluxes………………………………………………………………………………………………………………………151
List of Figures

Figure 1-1. A schematic description of the stomatal control on photosynthesis and transpiration of a leaf. .................................................................17

Figure 1-2. The spectra of vegetation and its variation with leaf chlorophyll content. .................................................................21

Figure 2-1. Schematic descriptions of the three upscaling schemes: BL, TBL and TL. .................................................................53

Figure 2-2. Simulated and measured annual ET at the studied sites, as well as the ET components obtained using BL, TBL and TL. .................................................................57

Figure 2-3. Simulated and measured annual GPP at the study sites, as well as the GPP components obtained using BL, TBL and TL. .................................................................68

Figure 2-4. The biases of annual ET and GPP estimates along with site mean LAI for each upscaling scheme. .................................................................59

Figure 2-5. The amount of light-saturated leaves at each site using BL and TBL (TL). .................................................................61

Figure 2-6. The difference between the ET estimates (a) and GPP estimates (b) from TBL and TL for sunlit leaves and shaded leaves. .................................................................62

Figure 2-7. Comparison between the average daytime $g_c$ obtained from TBL and the $g_s \times$ LAI obtained from TL for sunlit and shaded leaves. .................................................................64

Figure 2-8. Comparison between the daytime average $C_i$ obtained from TBL and TL for sunlit and shaded leaves. .................................................................65

Figure 2-9. A schematic description of the difference between TL and TBL models. .................................................................66

Figure 2-10. WUE for sunlit and shaded leaves estimated by TBL and TL. .................................................................67

Figure 3-1. Scatter plots of the measured leaf-level physiological parameters of C3 broadleaf leaves at the top of canopy at the Borden site in 2014 and 2015. .................................................................101

Figure 3-2. Seasonal variations of (a) Chlleaf, leaf area index (LAI), specific leaf area (SLA), and; (b) simulated $V_{\text{max},0}^{25}$ for each modelling scenario. .................................................................102
Figure 3-3. Scatter plots of the predicted and measured fluxes of 2013-2014 at the Borden site.................................................................103

Figure 3-4. The monthly relative MAEs (a, b) and the MAEs (c, d) between the estimated and measured daily fluxes in 2013-2014.................................105

Figure 3-5. Sensitivity of GPP and ET to the incoming solar radiation (Rs), air temperature (Tair), relative humidity (RH), LAI and Chlleaf at the Borden site……107

Figure 3-6. Seasonal patterns of the average daily incoming solar radiation (W/m²) on sunlit and shaded leaves simulated by BEPS in 2013 and 2014.........................108

Figure 3-7. BEPS modelled leaf photosynthesis and transpiration rates and their relationship to Chlleaf.................................................................109

Figure 3-8. Physiological differences, from BEPS simulations, between sunlit and shaded leaves.................................................................110

Figure 4-1. Distribution of the studied FLUXNET sites.................................132

Figure 4-2. Flow chart of the structure of BEPS-REF based on leaf area index (LAI) and clumping index (CI) and BEPS-LCC based LAI, CI and leaf chlorophyll content (LCC).................................................................135

Figure 4-3. Mean LCC, LAI and Vcmax25 for nine PFTs per week.................138

Figure 4-4. The reduction of biases and the improvement of R² for GPP and ET estimated by BEPS-LCC compared to the fluxes estimated by BEPS-LCC........141

Figure 4-5. The biases of estimated GPP along with the range of biases.........142

Figure 4-6. The biases of estimated ET along with the range of biases.........143

Figure 4-7. Daily variations of (a) GPP and (b) ET explained by variables including LCC, LAI, incoming solar radiation (Rs), air temperature (Tair), relative humidity (RH), precipitation(P) and wind speed (Ws) based on measurements from 124 flux towers.................................................................148

Figure 4-A1. Validation of annual GPP (a, b) and ET (c, d) from BEPS-LCC (a, c) and BEPS-REF (b, d)................................................................153

Figure 4-A2. Validation of site annual GPP for each PFT...............................154
Figure 4-A3. Validation of site annual ET for each PFT…………………………..155

Figure 4-A4. Uncertainties of the improvements brought by BEPS-LCC relative to BEPS-REF……………………………………………………………………………157

Figure 4-A5. Energy closure ratio of FLUXNET sites for each PFT………………158
Chapter 1

Introduction

1.1 Characteristics of terrestrial evapotranspiration

The earth has been going through a period of rapid changes since the beginning of the industrial era, due to a series of anthropogenic influences, especially the CO$_2$ emission from burning fossil fuels. These changes are displayed as unprecedented climatic conditions and consequent adjustments of the biosphere [IPCC, 2013]. Therefore, there is a strong demand for the science community to understand the trajectory and mechanism of climate change. Global land surface evapotranspiration (ET) is a critical flux in the global climate system, since ET accounts for a considerable amount of the terrestrial water budget and also dominates the energy balance on the land surface through turning liquid and solid water into gaseous forms.

ET is the collective upward flux of water molecules moving from the biosphere to the atmosphere. This water flux usually involves a phase change of water from liquid (or ice) to gas, which absorbs energy and cools the ambient environment. The latent heat accompanying ET is often referred to as $\lambda E$, where $\lambda$ is the latent heat for vaporizing a unit of water. ET is divided into evaporation and transpiration depending on the sources of outgoing water molecules [Scott and Biederman, 2017]. Water molecules vaporized from free water surfaces (i.e. lake, river and snow) and from soil constitute the flux of evaporation, which is a solely physical process determined by available energy, atmospheric water demands and soil properties. Transpiration refers to the flux of water molecules that move through the vascular plants and escape to the atmosphere through stomata on leaves. It is driven by a combination of physical environmental controls and restrained by plant physiology.

1.1.1 Global land surface ET

The quantification of annual ET, the long-term trend of ET and the partitioning of ET have drawn intensive attention in the context of climate change. In the past decade,
scientists have made significant progress in reducing the uncertainty in the range of terrestrial ET values via various data sources and techniques, including, but not limited to, isotope, eddy-covariance, runoff and precipitation inventory, remote sensing and process-based models. Table 1 compiles the terrestrial ET values reported since 2006 in comparable units.

**Table 1-1.** Global land surface evapotranspiration (ET). Numbers in bold are the values in the original units of each study. Latent heat λ is regarded as 2465 kJ/kg assuming global average air temperature is 15 °C.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Units (W/m²)</th>
<th>Units (mm/yr)</th>
<th>Units (10³×km³/yr)</th>
<th>Year</th>
<th>Land surface (10³× km²)</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Dirmeyer et al., 2006]</td>
<td>21.2-35.2</td>
<td>272-441</td>
<td>36.7-59.5</td>
<td>1986-1995</td>
<td>135</td>
<td>LSMs</td>
</tr>
<tr>
<td>[Oki and Kanae, 2006]</td>
<td>38</td>
<td>485</td>
<td><strong>65.5</strong></td>
<td>Nominal year</td>
<td>135</td>
<td>Various sources</td>
</tr>
<tr>
<td>[Fisher et al., 2008]</td>
<td>34-36</td>
<td>434-458</td>
<td><strong>57.8-61</strong></td>
<td>1986-1993</td>
<td>149</td>
<td>Remote sensing</td>
</tr>
<tr>
<td>[Wang and Liang, 2008]</td>
<td><strong>47.5</strong></td>
<td>608</td>
<td>82</td>
<td>1986-1995</td>
<td>135</td>
<td>Remote sensing</td>
</tr>
<tr>
<td>[Trenberth et al., 2009]</td>
<td><strong>38.5</strong></td>
<td>493</td>
<td>73.4</td>
<td>1985-1989</td>
<td>149</td>
<td>Various sources (top down)</td>
</tr>
<tr>
<td>[Jung et al., 2010]</td>
<td>43</td>
<td>556</td>
<td><strong>65±13</strong></td>
<td>1982-2008</td>
<td>117</td>
<td>Statistical model</td>
</tr>
<tr>
<td>[Yuan et al., 2010]</td>
<td>33</td>
<td><strong>417±38</strong></td>
<td>49</td>
<td>2000-2003</td>
<td>117</td>
<td>Remote sensing</td>
</tr>
<tr>
<td>[Haddeland et al., 2011]</td>
<td>32-46</td>
<td><strong>415-586</strong></td>
<td><strong>60-85</strong></td>
<td>1985-1999</td>
<td>~135</td>
<td>LSMs</td>
</tr>
<tr>
<td>[Lawrence et al., 2011]</td>
<td>38</td>
<td>489.1</td>
<td>66</td>
<td>2000</td>
<td>135</td>
<td>LSMs</td>
</tr>
<tr>
<td>[Mu et al., 2011]</td>
<td>34</td>
<td>432</td>
<td><strong>62.8</strong></td>
<td>2000-2006</td>
<td>117</td>
<td>Remote sensing</td>
</tr>
<tr>
<td>[Mueller et al., 2011]</td>
<td><strong>46±5</strong></td>
<td><strong>580±69</strong></td>
<td><strong>68±8</strong></td>
<td>1989-1995</td>
<td>117</td>
<td>LSMs</td>
</tr>
<tr>
<td>Reference</td>
<td>Range</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Year</td>
<td>Units</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------</td>
<td>------</td>
<td>--------------------</td>
<td>------</td>
<td>-------------------------</td>
<td></td>
</tr>
<tr>
<td>[Ryu et al., 2011]</td>
<td>39±8</td>
<td>500±104</td>
<td>63±13.1</td>
<td>2001-2003</td>
<td>133.7 TBM</td>
<td></td>
</tr>
<tr>
<td>[Jasechko et al., 2013]</td>
<td>39±5</td>
<td>499±64</td>
<td>73±9</td>
<td>Nominal year</td>
<td>Unknown Isotope</td>
<td></td>
</tr>
<tr>
<td>[Mueller et al., 2013]</td>
<td>38.5</td>
<td>493</td>
<td>64.6</td>
<td>1989-2005</td>
<td>131 TBMMs, LSMS, Reanalysis</td>
<td></td>
</tr>
<tr>
<td>[Wang-Erlandsson et al., 2014]</td>
<td>43</td>
<td>555</td>
<td>73.9</td>
<td>1999-2008</td>
<td>133.1 TBM</td>
<td></td>
</tr>
<tr>
<td>[Jiang and Ryu, 2016]</td>
<td>38±6.4</td>
<td>486±82</td>
<td>65±11</td>
<td>2001-2011</td>
<td>133.7 TBM</td>
<td></td>
</tr>
<tr>
<td>[Yao et al., 2016]</td>
<td>39.7</td>
<td>508</td>
<td>68.6</td>
<td>1970-2005</td>
<td>135 EC and LSMS</td>
<td></td>
</tr>
<tr>
<td>[Zhang et al., 2016b]</td>
<td>42.1</td>
<td>538.1</td>
<td>63.2</td>
<td>1981-2012</td>
<td>117 SIF and LSMS</td>
<td></td>
</tr>
</tbody>
</table>

Among the studies listed in Table 1-1, midpoint estimates of terrestrial ET are distributed in a range from 50 to 75 × 10^3 km^3/yr, equivalent to 450 to 600 mm/yr or 30 to 45 W/m^2 (but with a full uncertainty spanning 22-55 W/m^2). One major source of uncertainty stems from the difference among the land surface areas used in these studies. Although the global land area is about 149 million km^2, most ET products tend to omit the continent of Antarctica [Lawrence et al., 2011; Mueller et al., 2013] and, in some cases, also exclude Greenland [Haddeland et al., 2011], resulting in a land surface area of around 130-135 million km^2. The omission of cold areas does not influence the ET value in km^3/yr because sublimation contributes only a small fraction of the outgoing water flux, but the area-dependent units (i.e. mm/yr and W/m^2) are substantially affected. Some ET products only focus on vegetated areas (117 million km^2) because ET contributed from perennial deserts is minimal, and these products generally show a higher ET in units of mm/yr and W/m^2 than other studies. Considering the complexity brought by units, several studies [Jung et al., 2010; Ryu et al., 2011] have strongly suggested that km^3/yr is a better unit to report global ET since it is relatively insensitive to the exclusion of low ET areas.
In addition to the unit, each method is affected by the uncertainty underlying its assumptions. Before 2006 there is a lack of study on global terrestrial ET. Simulations from multiple LSMs in the International Satellite Land Surface Climatology Project (ISLSCP) provide a rough guess of terrestrial ET with large discrepancy among models [Dirmeyer et al., 2006]. Meanwhile, many global ET maps were quickly produced using remote sensing data in the framework of simplified the Penman-Monteith equation and Priestley-Taylor equation [Cleugh et al., 2007; Fisher et al., 2008; Wang and Liang, 2008]. Though empirical methods have proven to be reliable at some sites, they often suffer from limited locality and are known to incur errors in regions beyond their validation spectrum [Wang and Liang, 2008]. Jung et al. [2010] reported the terrestrial ET is around 65±3 × 10³ km³/yr, by assimilating ET measurements from global eddy-covariance towers into a neural network and using explanatory climatic values to extrapolate ET to the continuous land surface. However, the bottom-up method estimates ET smaller than a top-down method, which derived ET as the residue of precipitation and runoff [Trenberth et al., 2009]. The difference between the two method is at the order of ~10 × 10³ km³/yr, but the whereabouts of the “dark” ET is yet to be explored.

The studies in the second half of the last decade illustrate the ET estimates from process-based models, either from prognostic land surface models (LSMs) or from diagnostic terrestrial biosphere models (TBMs). ET estimates from LSMs vary over a wide range [Mueller et al., 2013], but LSMs assimilated with solar-induced chlorophyll fluorescence (SIF) [Zhang et al., 2016b] or eddy-covariance measurements [Yao et al., 2016] can produce results that are rather close to the statistical values reported by Jung et al. [2010]. TBMs generally produce ET values close to the statistical results, because they normally use satellite observations to constrain the terrestrial vegetation dynamics [Liu et al., 2003; Ryu et al., 2011] and consider more biophysical processes than LSMs [Fisher et al., 2014]. The necessity to use external data sources, especially those representing vegetation carbon uptake, as a constraint in process-based models, indicates potential unknown processes or uncertain parameters in ET models.

By assembling the ET estimates from the reliable statistical model and the LSMs constrained by external datasets, the best estimate of ET is around 38-40 W/m², 500-550
mm/yr and $65-70 \times 10^3 \text{ km}^3/\text{yr}$ for the global land surface without Antarctica. This ET value accounts for two thirds of the $110 \times 10^3 \text{ km}^3/\text{yr}$ precipitation that falls over the landmass [Oki and Kanae, 2006; Jasechko et al., 2013] and consumes about 50% of net solar radiation absorbed (70-80 W/m$^2$) by the land surface [Trenberth et al., 2009].

1.1.2 Trends and variability of ET

With the consistently rising global surface temperature, the global hydrological cycle, including the land surface ET branch, is expected to intensify. However, there has been a discrepancy over the direction and magnitude of the ET trend, and the attributions of the trend.

At the global scale, terrestrial ET seems to have experienced an increase from the early 1980s to 2000 [Wild et al., 2008; Jung et al., 2010; Wang et al., 2010a]. Wang et al. [2010] found that global evapotranspiration increased at a rate of 0.06 W/m$^2$/yr from 1982 to 2002. Wild et al. [2008] found a smaller increasing trend at 0.019 W/m$^2$/yr from 1982 to 2000, but the rate intensified to 0.05 W/m$^2$/yr after 1992. The enhancement of ET is attributed to a decade of global brightening and the warming effect of increasing atmospheric CO$_2$ concentration. However, Jung et al. [2010] found a turning point around 1998, where ET plateaued or even declined afterward, coinciding with a declining trend of soil moisture corroborated by satellites. Miralles et al. [2013] confirmed the ET decline in the 21st century but attributed it to the variability in the climatic system (i.e. El Niño–La Niña cycle). Several recent studies, however, reported a sustained increasing trend [Zhang et al., 2015; Yao et al., 2016]. Yao et al. [2016] reported that terrestrial ET increased by 0.018 W/m$^2$/yr (p<0.05) from 1970-2005 by synthesising ground based eddy-covariance data in CMIP5 LSMs. Zhang et al. [2015] reported an increasing trend in ET during 1981 to 2012 at the rate of 0.044 W/m$^2$/yr (p<0.01), and suggest that this increase was due to enhanced transpiration through global greening, which is indicated by increasing NDVI, and due to increasing air vapor pressure deficit. However, one may be concerned with the credibility of the detected ET trends (<0.06 W/m$^2$/yr) since they are relatively small with respect to the uncertainty of modelled global ET (~5 W/m$^2$) noted in Section 1.1.1. The uncertainty of ET reflects either the inter-annual variations of ET or the uncertainty of inputs used in modelling ET.
To analyse the effect of inter-annual variations on the trend detection, a study drew 1000 30-year gross primary productivity (GPP) series from random GPP values generated based on an uncertainty of 4 Pg/yr under a Gaussian distribution, fitted a linear trend to each of these sampled series and found that a trend should be larger than 0.0053 Pg/yr² to exceed the 95% confidence interval of the randomly sampled trends [Baldocchi et al., 2016]. A similar study is not available for ET, though the trends of ET from literature seem to be beyond 95% confidence interval after comparing the relative magnitude of the trends to the uncertainty for GPP and that for ET. However, the ET uncertainty propagated from its critical modelling inputs is still elusive, such as those of LAI reported by a recent study [Jiang et al., 2017].

There are also strong spatial variations in the ET trend. The continents above 30 N witnessed the most significant increase in ET, while the tropical areas and southern hemisphere (i.e. Australia, South Africa) show no increase, or even decreasing ET trends [Jung et al., 2010; Douville et al., 2012; Miralles et al., 2013]. However, this pattern is at odds with the results from CMIP5 modelling studies, which show consistent increasing ET across the globe [Yao et al., 2016], and one study also finds an increasing trend in Australia [Zhang et al., 2015]. Wang and Dickinson [2012] also found that the differences among models at a regional scale are much larger than those of the global averaged values. These discrepancies among studies and the spatial variations in the ET trend signify the necessity to understand and quantify the inter-annual processes driving the variability of ET.

1.1.3 Transpiration and water use efficiency

Transpiration describes the water usage of plants. It is a key process in the interaction between ecosystems and climate. Although transpiration (T) is known to be a major component of global terrestrial ET, the amount of T, or the T/ET ratio, remains far from certain [Coenders-Gerrits et al., 2014; Wei et al., 2017]. Table 2 summarizes recent studies on global average T/ET ratios.
Table 1-2. Components of global terrestrial ET. T refers to transpiration, Ei refers to evaporation of intercepted rainfall and snow on canopies, Es refers to the evaporation from wet soil surface and inland water surface.

<table>
<thead>
<tr>
<th>Reference</th>
<th>T/ET (%)</th>
<th>Ei/ET (%)</th>
<th>Es/ET (%)</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Dirmeyer et al., 2006]</td>
<td>48</td>
<td>16</td>
<td>36</td>
<td>model</td>
</tr>
<tr>
<td>[Lawrence et al., 2011]</td>
<td>56</td>
<td>28</td>
<td>31</td>
<td>model</td>
</tr>
<tr>
<td>[Miralles et al., 2011]</td>
<td>80</td>
<td>11</td>
<td>9</td>
<td>remote sensing</td>
</tr>
<tr>
<td>[Jasechko et al., 2013]</td>
<td>80 - 90</td>
<td></td>
<td></td>
<td>isotope</td>
</tr>
<tr>
<td>[Coenders-Gerrits et al., 2014]</td>
<td>35 - 80</td>
<td></td>
<td></td>
<td>isotope</td>
</tr>
<tr>
<td>[Schlesinger and Jasechko, 2014]</td>
<td>61 ± 15</td>
<td></td>
<td></td>
<td>literature review</td>
</tr>
<tr>
<td>[Wang-Erlandsson et al., 2014]</td>
<td>59</td>
<td>21</td>
<td>20</td>
<td>model</td>
</tr>
<tr>
<td>[Good et al., 2015]</td>
<td>64 ± 13</td>
<td></td>
<td></td>
<td>isotope</td>
</tr>
<tr>
<td>[Zhang et al., 2016b]</td>
<td>65</td>
<td>10</td>
<td>25</td>
<td>model</td>
</tr>
<tr>
<td>[Wei et al., 2017]</td>
<td>57.2 (42.9 - 74.9)</td>
<td></td>
<td></td>
<td>remote sensing</td>
</tr>
<tr>
<td>[Fatichi and Pappas, 2017]</td>
<td>70 ± 9</td>
<td></td>
<td></td>
<td>model</td>
</tr>
</tbody>
</table>

Table 1-2 shows that isotopic measurements and models are the two major methods to obtain the T/ET ratio, but their reported values are distributed over a wide range (from 50% to 90%), with recently reported values likely to be constrained between 55-80%. The majority of the uncertainty in the T/ET ratio is propagated from the limitations of each method. The isotope technique is the basis of the only observational effort at large scales to separate T from ET. The ratio of heavy to light isotopes of water (i.e. D/H) in transpiration is often assumed to be unchanged relative to soil water, while evaporation will have higher D/H because of the lower vapor pressure and lower diffusivity of the heavy deuterium isotope [Good et al., 2015]. By measuring and comparing D/H in precipitation, runoff and ET, it is possible to partition ET into evaporation and
transpiration. However, the results are very sensitive to the geographic distribution of basins and flux towers used for the analysis [Coenders-Gerrits et al., 2014; Good et al., 2015]. The simple empirical models using remote sensing products to estimate T and E separately show different T/ET ratios from study to study, meaning that a simple relationship between T/ET and LAI or other indices might not be applicable on a global scale [Miralles et al., 2011; Wei et al., 2017]. LSMs consider the physical processes in the soil-vegetation-atmosphere continuum and suggest that the T/ET ratio is around 60% on average, while there remains 20-30% variability among model results due to differences in model structure and parameterization [Lawrence et al., 2011; Wang-Erlandsson et al., 2014].

The T/ET ratio is also found to vary with vegetation status and plant functional type. Though limited studies have compared them, they show that cropland has the highest T/ET ratio, from 60 to 90%, while natural vegetation usually shows T/ET ratios of around 50 to 70% [Wei et al., 2017]. Baldocchi and Vogel [1996, 1997] studied the ET from the overstories and understories at a temperature broadleaf forest and a boreal jack pine forest. They found that the evaporation from forest floor only accounts for 5% of peak ET for broadleaf forest, while this ratio is about 10-40% for the jack pine forest. Through a synthesis of T/ET studies, Schlesinger and Jasechko [2014] find that forests show higher T/ET than shrubland and grassland, and tropical vegetation has higher T/ET than vegetation in temperate and boreal regions. It seems the PFT with dense canopies tend to have a higher T/ET ratio.

The trends of transpiration are often investigated in combination with carbon uptake by vegetation, because the exchanges of CO₂ and H₂O are coupled in that they are both controlled by stomata – the small pores on the leaf surface. Water use efficiency (WUE) has long been used as the metric to describe the trade-off between water loss from and carbon gain by plants. Though the definition of WUE varies with scales and disciplines [Bernacchi and VanLoocke, 2015], the intrinsic WUE (iWUE), which removes the effects of VPD variability on WUE, is a commonly adapted indicator to evaluate leaf and ecosystem functions [Keenan et al., 2013]. A rising iWUE has been demonstrated by many studies at different scales [Zhou et al., 2013; Bernacchi and VanLoocke, 2015] but
is attributed to different drivers. Through testing many competing hypotheses, Keenan et al. [2013] suggested that the increase in iWUE is driven by rising atmospheric CO$_2$ concentrations, but current biosphere models predict this effect to be much smaller than inferred. Holmes [2014] suggested that high concentrations of ozone and other phytotoxic air pollutants injure the trees and reduce their photosynthesis more than transpiration. Therefore, the reduced air pollution in past two decades may explain part of the increasing trend of iWUE. Tang et al. [2015] also added the land use and cover as another driver, suggesting that different vegetation types have different iWUE – forests show higher WUE than crops and grasses.

The increase in iWUE can be interpreted as either transpiration decrease or photosynthesis increase or both occurring simultaneously. Most evidence points to an increasing carbon sink, explained by the CO$_2$ fertilization effect, with a possible contribution from decreasing ET [Keenan et al., 2013]. However, a recent isotope-based study showed that the increase in iWUE has no apparent effect on tree growth, suggesting that changes in transpiration must be stronger than that of photosynthesis [Silva and Horwath, 2013]. Frank et al. [2015] found a simultaneous increase in iWUE and transpiration, and attributed the increasing transpiration to longer growing seasons and higher LAI. According to modelling results forced by an increasing WUE with CO$_2$, Knauer et al. [2017] found that modelled ET, carbon exchange and runoff conflict with the observations from various sources for forest ecosystems. This result suggests that the WUE values derived from models are considerably larger than real ecosystem WUE and that WUE is not sensitive to CO$_2$.

Over all, in the past decade we have made significant progress in quantifying the global terrestrial ET. Meanwhile, the trend, spatial and temporal variability, and partitioning of ET are open to debate, and definitive explanations for their apparent changes remain elusive. Globally, transpiration accounts for at least half of total ET, but its trend and how it interacts with climatic controls and biochemical processes in plants are unclear. More observational effort is needed to find the major controls for ET, especially T, and to incorporate the resulting knowledge into process-based models.
1.2 Measurement and modelling of ET

1.2.1 Eddy covariance technique

Since the mid-1990s, the eddy-covariance (EC) technique has become a mainstream tool to directly measure the carbon and water fluxes between the biosphere and the atmosphere [Baldocchi et al., 2001]. The EC technique uses high frequency measurements (10-20 Hz) of wind speed and wind direction as well as the CO$_2$ and H$_2$O concentration in the air at a point above the canopy using a three-axis anemometer and a fast response infrared gas analyzer. If the vertical wind speed is $w$ (m/s) and the specific humidity of the air mass is $q$ (g/kg), $\rho_a$ is the density of air (kg/m$^3$), then the water flux density is defined as the covariance of $w$ and $q$:

$$E = \frac{1}{N} \sum_{i=1}^{N} (w_i - \bar{w})(q_i - \bar{q}) \rho_a$$  

where $\bar{w}$ is the mean vertical wind speed and $\bar{q}$ is the average specific humidity in the air, $N$ is the number of records in a half-hourly or an hourly step. A positive flux density means water vapor is transferred into the atmosphere and away from the land surface and negative values denote the reverse [Baldocchi et al., 2001]. High frequency measurements are typically integrated over periods of half an hour [Goulden et al., 1996], building the basis to calculate carbon and water exchanges from daily to annual scales. Since horizontal advection can also affect vapor density on patchy and uneven surfaces, it is important to place the flux tower – the scaffold structure whose height is usually double the canopy height and where EC instruments are mounted – in the middle of a flat and homogeneous land surface.

Currently, a global network comprised of more than 400 active flux towers is providing continuous, long term, high temporal resolution ET measurements across a wide range of plant functional types (http://fluxnet.fluxdata.org/). A set of protocols, including partitioning the carbon fluxes into gross primary productivity (GPP) and ecosystem respiration (ER) [Reichstein et al., 2005; Lasslop et al., 2010], filling data gaps in carbon and water fluxes [Barr et al., 2004] and constraining the random uncertainties of measurements (i.e. placement of tower, instrument) [Richardson et al., 2006], have been
developed to process flux data. The latest effort to standardize flux measurements from across various regional flux networks (e.g. AmeriFlux, AsiaFlux, OzFlux) resulted in the FLUXNET2015 dataset [Papale et al., 2006].

One key problem involved in the ET measurements from EC is the energy balance problem. In theory, net radiation (Rn), soil heat flux (G), sensible heat (H) and λE measured at a site should converge to the energy balance: Rn – G = λE + H. However, the sum of λE and H sensed by EC instruments is often smaller than the difference of Rn – G, suggesting that ET measurements from flux towers are underestimated. Past reviews found the average hourly energy balance closure ratio of (λE + H) to (Rn – G) to be around 80% at many EC sites [Wilson et al., 2002; Wang and Dickinson, 2012; Stoy et al., 2013]. According to the study by Stoy et al. [2013], EC measurements at evergreen forests and savannas display the best energy closure ratio (0.91-0.94), while crop, deciduous forest and wetland show the lowest ratio (0.70-0.78).

Many theories have been proposed to explain the energy non-closure of high frequency EC measurements. These theories include the mismatch of instrument footprints, observational errors in some energy components and violation of the EC technique assumptions due to advective flux divergence [Gao et al., 2017]. However, the leading hypotheses are currently the heat storage theory [Leuning et al., 2012] and the big eddies theory [Foken, 2008].

Leuning et al. [2012] suggested that a substantial part of the underestimation in ET can be explained by phase lags between the measured fluxes and the energy stored in soils, air and biomass below the measurement height. When using half-hourly averages fluxes, energy closure is only observed at 8% of flux sites in the La Thiuile dataset – the predecessor of the FLUXNET2015 dataset, but energy closure is realised at 45% of the sites when using 24-hour averages because energy entering the soil and biomass in the morning is returned in the afternoon and evening. By meticulously measuring the energy storage component, closure was achieved for cropland [Meyers and Hollinger, 2004] and forest [Lindroth et al., 2010].
Foken et al. [2008, 2011] reviewed the results of multiple large surface flux campaigns and argued that the problem of eddy covariance energy balance closure is fundamentally a problem of scale: the missing of low frequency big eddies. Big eddies that flow at a scale comparable to the tower footprint usually do not touch the ground and thus cannot be sensed by EC equipment very often. Kanda et al. [2004] suggests that the finite (half-hourly) time averaging of EC measurements would randomly include and omit the energy contributed by large eddies because of their less frequent appearance, and the phase difference between the vertical velocity and scalar of interest (i.e. CO₂ and H₂O concentration) in large eddies contribute the most to the energy non-closure. A recent study found that the enlarged phase difference in big eddies can explain the degree of ET underestimation, because the big eddies are often linked to prevailing advection occurrences [Gao et al., 2017].

To correct the energy closure problem, Twine et al. [2000] proposed a method that assumes that the Bowen Ratio – the ratio of sensible heat to latent heat – is preserved during a period of half an hour or an hour and can be used to divide the missing energy into sensible heat and latent heat. This method has been adopted by many studies and data producers [Wohlfahrt et al., 2009; Jung et al., 2010]. In addition, using large time-scales to average out the temporal lag in storage would provide ET measurements with a higher closure ratio [Leuning et al., 2012], assuming the Rₙ is homogeneous over the EC footprint which is rarely the case.

However, a concern still exists on whether the energy closure problem is indeed a problem? Scott [2010] compared the eddy covariance measurements at a grassland, a shrubland and a savanna site with their small catchment water balance. ET from these two datasets agree within 3% of the annual values. An earlier study also found annual ET from EC is similar to the ET from catchment water balance at a mixed forest [Wilson et al., 2001]. Daily EC measurements significantly covary with sap flow and soil water budget, which are independent indicators of transpiration and ET. It is suggested that the closure problem may simply caused by the phase difference or the footprint different of the energy components [Baldocchi and Vogel, 1997] other than latent heat and sensible heat. These results give us confidence in using EC data to evaluate model performance in
Some studies use EC measurements to improve ET models through model-data fusion [Williams et al., 2009]. In this thesis, EC measurements will be the major source of data for assessing model performance.

1.2.2 Modelling of ET

The variability in ET is driven by a diverse range of biotic and abiotic factors (i.e. available energy, precipitation and humidity). A review of 46 site-level studies suggests that vegetation predominantly controls the inter-annual variability of ET and the performance of process-based ET models is highly dependent on the description of vegetation status [Wang and Dickinson, 2012]. The influence of abiotic factors on ET varies from site to site: some case studies found that in tropical areas, solar radiation is the strongest determinant of ET [Fisher et al., 2009; Costa et al., 2010], while precipitation and deep soil water are more relevant to ET in semi-arid areas [Leuning et al., 2005], and ET of boreal regions is affected by both available energy and VPD [Kelliher et al., 1997]. Building upon the knowledge obtained from site-level studies, a hierarchy of semi-empirical models were developed to simulate ET, utilizing equations that have concise physiological meaning and require only a few input variables [Running et al., 2004; Fisher et al., 2008; Wang et al., 2010b; Mu et al., 2011; Yebra et al., 2013]. These models particularly rely on remotely-sensed data to account for the factor of vegetation. They provide quick and valuable references for the large-scale mapping of ET. However, some cross-site validations have shown considerable biases in these products (e.g. MODIS) due to their over-simplified view of the biophysical and biochemical processes in the biosphere [Heinsch et al., 2006; Zhang et al., 2012; Jiang and Ryu, 2016]. In order to obtain accurate estimates of global ET, a genre of advanced statistical models that use a machine learning approach to upscaling flux measurements with gridded explanatory variables to regional scales was developed [Beer et al., 2010; Jung et al., 2010]. The machine learning approach relies on the prior knowledge of flux measurements; it is known to provide the most credible estimates so far but also lacks descriptions of the land surface processes and predictive abilities. Nowadays, the process-based terrestrial biosphere models (TBM) that consider the key land surface
processes are the principle tools used to simulate global ET, explore the underlying mechanisms of ET processes, and investigate the response of ET to environmental changes.

The widely-used Penman-Monteith (PM) equation provides a process-based framework for TBMs to simulate ET. It incorporates all the biotic and abiotic factors in an equation derived based on physical laws [Campbell and Norman, 2012]. Considering vegetation on the land surface to be analogous to a simple big-leaf, PM aggregates all biotic controls on ET into a canopy conductance \( (g_c) \) to describe the resistance on the carbon and water exchanges between plants and ambient air. The PM equation is expressed in the following form:

\[
ET = \frac{\Delta (R_n - G) + \rho c_p VPD g_v}{\Delta + (1 + \frac{\rho c_p}{g_v}) \gamma} \times \frac{1}{\lambda}
\]  

(1-2)

where \( \lambda \) is the latent heat of water, \( R_n \) is the net radiation at the leaf surface, \( G \) is the soil heat flux, \( \rho \) is the density of air, \( c_p \) is the specific heat of air, \( VPD \) is the vapour pressure deficit of the ambient air, \( \gamma \) is the psychrometric constant, \( g_v \) is leaf boundary layer conductance for water vapour, \( \Delta \) is the slope of the saturation vapour pressure curve at air temperature, and \( G_c \) is the canopy conductance.

\( G_c \) is the reciprocal of canopy resistance \( (r_c) \), which is an analog to the resistance for electricity to define the controls of vegetation, soil and boundary layers on diffusions of CO\(_2\) and H\(_2\)O through canopies. Resistances can work in parallel or in series to control the flows, and the total resistance is calculated in a similar way to electricity study. In this thesis, conductance is a more convenient concept, considering that leaves work in parallel to exchange CO\(_2\) and H\(_2\)O with the atmosphere. Most TBMs apply a two-layer frame, in which the ET from vegetation and the evaporation from soil are considered separately [Hanson et al., 2004]. A canopy stomatal conductance \( (G_s) \) is used to describe the control of vegetation. \( G_s \) accounts for majority of \( G_c \), though they are not equivalent to each other [Baldocchi and Meyers, 1998, Kelliher et al., 1995]. When a canopy is closed and evaporation from soil contributes little to total ET, the value of \( G_s \) is close to that of \( G_c \). Other variables used in the PM equation are simulated and provided by their
corresponding components in TBMs, such as the radiative transfer module, soil thermal module and biochemical process module.

The modelling community has relentlessly improved TBMs and reported promising ET simulations at the site-level compared to EC measurements [Zaehle and Friend, 2010; Zhang et al., 2016a]. However, TBMs still face considerable uncertainties when extrapolating to other ecoregions because of the difficulty in parameterizing key vegetation characteristics at large scales [Zaehle et al., 2005; Groenendijk et al., 2011]. Moreover, the correlation between ET estimates and EC measurements is often limited with r² values around 0.7 for the TBMs that provide global simulations [Ryu et al., 2011; Jiang and Ryu, 2016]. Since vegetation is the dominant control on ET and G_s is the only parameter in the PM equation representing the biotic regulation, a further attempt to improve ET estimations leads to an obvious question of improving the estimates of G_s.

The Penman-Monteith equation combines physical and physiological constraints into one simple equation for ET estimations. The physical constraints are determined by the available energy that provides the latent heat of vaporization and the atmospheric humidity and wind that cause the drying power of the atmosphere, while the physiological constraint is described using the canopy conductance. However, the simplicity of the Penman-Monteith equation also leads to a potential imperfection: the Penman-Monteith equation is often applied to a plant canopy with a big leaf assumption that the physiological constraint of the canopy can be scaled from leaf to canopy using the canopy conductance. This assumption is apparently in conflict with the complex structure of plant canopies in reality, where the leaf distribution varies by clumping [Chen et al., 1997], light environments [Norman et al., 1982] and angle and heights [Baldocchi and Meyers, 1988] and consequently influences the canopy transpiration rates. However, many studies have shown the reliability of the Penman-Monteith equation for ET estimation regardless of this defect, mainly because they treat ET as an independent process, and G_c or G_s can be freely tuned to fit ET measurements. In the late 1980s, some ecological studies have found that water loss from leaves is tightly coupled with carbon gain, meaning that the rate of photosynthesis becomes a constraint on stomatal conductance and ET [Ball et al., 1987]. This water-carbon coupling principle
projects doubts on the previous ET studies that are simply dependent on the Penman-Monteith equation, because the tuning of $G_c$ and $G_s$ to fit ET measurements may not fit carbon flux measurements simultaneously. The canopy conductance for carbon cannot be derived from carbon flux measurements because it has been demonstrated that the big-leaf model for photosynthesis is theoretically flawed and inaccurate [Chen et al., 1999; Sprintsin et al. 2012] This issue of the Penman-Monteith equation has not been well recognized, and the Chapter 2 of this thesis will strive to address this issue.

Except the frame provided by the Penman-Monteith equation to directly simulate ET, another genre of ET models chooses to estimate other components of the energy budget (i.e. sensible heat) and calculates ET as the residue of balance. For example, in the Atmosphere-Land Surface Exchange Inverse (ALEXI) model [Norman et al., 1995; Anderson et al., 2008, 2011], thermal infrared remote sensing archives are integrated to diagnose the canopy temperature and soil temperature, then the temperature gradient between the surface and the atmosphere is utilized to calculate the sensible heat flux and consequently the latent heat flux. This genre of models provides a simple solution for regional ET modelling because it minimizes the need for ancillary meteorological data and avoids the complicated upscaling process from leaf to canopy.

1.2.3. Modelling stomatal controls

Theoretically, $G_s$ is equal to the reciprocal of the resistance collectively enforced by stomata on leaves, because leaf stomata work in parallel to transport water molecules. The openness of stomata is quantified by stomatal conductance ($g_s$) or its reciprocal – stomatal resistance ($r_s$). $g_s$ is found to be sensitive to a number of climatic factors, including air temperature, solar radiation and humidity. Empirical stomatal models assume that there is a maximum $g_s$ ($g_{smax}$) for leaves under an ideal environment, and that the real-time $g_s$ can be reduced from $g_{smax}$ using a series of scaling functions for various climatic factors [Jarvis, 1976; Stewart, 1988; Kelliher et al., 1995; Samson and Lemeur, 2001].
Progress in plant physiology since the 1970s revealed that the leaf photosynthetic rate \( (A_n) \) is the dominant factor in controlling \( g_s \) \cite{Wong et al., 1979, 1985}. Plants adjust \( g_s \) to curb CO\(_2\) absorption to maintain an optimal intercellular CO\(_2\) concentration in chloroplasts and an optimal water use efficiency \cite{Sellers, 1997}. \textit{Ball et al.} \cite{Ball et al., 1987} developed a semi-empirical stomatal conductance model (denoted “BWB stomatal model”) in which \( g_s \) is linearly related to \( A_n \), relative humidity (RH) and the reciprocal of atmospheric CO\(_2\) concentration. The influence from temperature, solar radiation and soil moisture on \( g_s \) were factored into the simulation of \( A_n \).

\[ \text{Figure 1-1. A schematic description of the stomatal control on photosynthesis and transpiration of a leaf. The exchange of CO}_2\text{ and H}_2\text{O between the air and the leaf is regulated by g}_s\text{. CO}_2\text{ molecules diffuse to chloroplasts of leaves to participate in photosynthesis driven by pigment (i.e. chlorophyll) and enzyme (i.e. Rubisco). The rate of photosynthesis (A}_n\text{) in turn controls g}_s\text{ and transpiration.} \]

The BWB stomatal model is a prevailing framework to describe the behaviour of stomata in LSMs and TBMs, such as the renowned Community Land Surface model (CLM). \textit{Leuning} \cite{Leuning, 1995} developed a modified version of the BWB model by replacing the linear dependence on RH with a hyperbolic dependence on the leaf-to-air vapor pressure deficit (VPD) and adding the CO\(_2\) compensation point to the model. The modified model is found to fit measurements of several species better than BWB models but it requires more empirically fitted variables. \textit{Medlyn et al.} \cite{Medlyn et al., 2011} unified the BWB and Leuning models in a process-based form, based on an optimality theory that stomata should act to
maximize carbon gains while minimizing water loss. Using this new model, Lin et al. [2015] found that $g_s$ differs among PFTs according to the reactions of its key parameter $g_1$ (a parameter similar to the reciprocal of WUE) to soil moisture, temperature, and hydraulic features of trees. This new stomatal model provides a framework to analyse the stomatal behaviour at large scales. However, since this framework is relatively new and much remains unknown about the behaviour of its key parameter $g_1$, this study still uses the widely-used BWB stomatal model in the TBM used in this study (Section 1.4.2).

In the BWB stomatal model, transpiration can be thought of as a “cost” or a “penalty” incurred by plants when they open their stomata to satisfy their photosynthetic demand for atmospheric carbon dioxide [Katul et al., 2012]. The BWB model forms the cornerstone to couple the water and carbon fluxes in process-based models. $A_n$ is a key variable in determining $g_s$, and thereafter in determining $g_c$ and ET.

Vegetation is the predominant factor that influences the terrestrial ET through controlling leaf stomatal conductance. The openness of leaf stomata is linearly regulated by photosynthetic rates, which are in turn driven by the biochemical processes in leaves. A strong representation of the vegetation structural and physiological traits and leaf biochemical processes in TBMs is the key to accurately simulating terrestrial ET.

1.3 Remote sensing of vegetation traits

1.3.1. Vegetation structural and physiological traits

The status of vegetation is described by a series of traits from the leaf-level to the canopy-scale, and generally classified into structural traits or physiological traits [Pallardy and Kozlowski, 2008]. Structural traits and physiological traits are often regarded as biophysical properties and biochemical properties, respectively, emphasizing their different functions in ecosystems [Asner, 1998].

Structural traits or biophysical properties describe the morphology, size and distribution of vegetation elements like leaves and stems in a canopy. Commonly measured traits include leaf area, woody area, leaf and stem distribution, foliage clumping, tree height
and understorey density. With an accurate description of vegetation structure in a canopy, it is possible to simulate the radiation [Ross, 1981], water and heat [Norman, 1982] transfer between the components, thereby elucidating the micrometeorological conditions inside the canopy that drive the biochemical processes (i.e. photosynthesis, respiration) of leaves. Conversely, structural traits can be used to upscale the results of leaf-level biochemical models to the ecosystem scale in LSMs and TBMs, in order to facilitate large-scale applications [Leuning et al., 1995; Chen et al., 1999]. In addition, these structural traits are important to determining the biomass of the plants [Friedl et al., 1994; Gitelson et al., 2003].

Plant physiological traits usually describe the content of pigments, enzymes, nutrients and compounds that support the function of vegetation, especially photosynthesis. Several biochemical processes including the Calvin-Benson cycle, photorespiration, mitochondrial respiration, light inhibition, non-photochemical quenching and fluorescence, jointly determine the rate of carbon uptake by ecosystems [Porcar-Castell et al., 2014]. Leaf chlorophyll, leaf nitrogen, and the photosynthetic enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) are the major physiological traits involved in carbon assimilation. The development of physiological traits is found to vary with phenology, species, nutrient availability and climatic factors [Xu and Baldocchi, 2003; Croft et al., 2017], while the variations of these traits are still not fully understood at the global scale.

The description of vegetation status is critical in TBMs. TBMs tend to use synoptic descriptions of physical processes on the land surface to achieve a balance between model efficiency and accuracy [Sellers, 1997]. Therefore, unessential plant traits such as wood density, and lignin and cellulose fractions are usually not considered in the parameterization of TBMs. For those traits that are included in TBMs, they are often simplified into a fixed value when we lack a way to utilize their limited ground measurements for large regions. LAI is by far the structural parameter with the strongest seasonality used in TBMs, and it can be obtained using remotely-sensed surface reflectance. In contrast, other key structural and physiological traits are usually assigned
according to ecoregion types or plant functional types, based on a limited number of measurements [Ustin and Gamon, 2010; Wullschleger et al., 2014].

It has been recognized that the relative underrepresentation of vegetation status in models is a major source of uncertainty for TBM estimations of ET. The lack of structural parameters often results in a false assumption of the micrometeorological conditions inside canopies. One major caveat of TBMs is to simplify a canopy as a big-leaf, though it has been recognized to cause considerable errors in GPP estimates [Sprintsin et al., 2012] and ET estimates [Baldocchi and Harley, 1995] at some sites. The big-leaf scheme does not consider foliage clumping and the consequent shaded environment for some leaves, in which the changing light environment significantly dampens photosynthesis. As for the representation of physiological traits in TBMs, the major goal is to better capture Vcmax25, which defines the maximum leaf photosynthetic capacity, and is regarded as the most critical physiological trait in TBMs [Rogers, 2014]. Vcmax25 can only be inferred from the measurement of A-Ci curves, through measuring the response of photosynthetic rate to changes in CO₂ concentration [Xu and Baldocchi, 2003]. It is a very time consuming procedure; thus, in modelling practice a constant value is often assigned to Vcmax25 according to the plant functional type [Wullschleger, 1993]. The incorporation of temporally and spatially varying Vcmax25 is regarded as a key pathway to upgrading TBMs [Wullschleger et al., 2014; Ali et al., 2015].

1.3.2 Remote sensing of physiological traits

Remote sensing holds a unique and critical role in providing large-scale observations of the land surface at regular intervals. Detecting vegetation status is a major application of remote sensing, since the selective absorption and reflectance of plants will allow sensors to gather a tremendous amount of information relevant to the structure and function of vegetation [Ollinger, 2011]. Our ability to derive the structural traits and physiological traits from satellites depends on our understanding of the ways that vegetation traits could affect the spectroscopy of leaves and canopies.

The basic form of vegetation reflectance is shown in Figure 2. The reflectance in the visible band (VIS; 400-750 nm) is typically low due to pigment absorption, and the
reflectance in the middle infrared (MIR; >1400 nm) is low due to water absorption. The near infrared region (NIR; 750-1400 nm) forms a high reflectance plateau which results from leaf and canopy structure-induced scattering.

![Figure 1-2](image)

**Figure 1-2.** The spectra of vegetation and its variation with leaf chlorophyll content. Reproduced from Ollinger [2011].

Many vegetation indices (VIs) exploit the NIR plateau to explore vegetation dynamics; to name a few: normalized difference vegetation index (NDVI) [Tucker, 1979], enhanced vegetation index (EVI) [Gao et al., 2000] and soil-adjusted vegetation index (SAVI) [Huete, 1988]. VIs are directly used to study ecosystem productivity [e.g. Myneni et al., 1997], biomass [e.g. van der Meer et al., 2001] and phenology [e.g. White et al., 1997] of the land surface, though their scalars lack a physical meaning. Since the NIR plateau mainly reflects the status of vegetation structure, VIs are widely used to derive a vegetation structural trait called leaf area index (LAI), which represents the area of the green surface supporting photosynthesis and transpiration in process-based TBMs. LAI is defined as one-half of the total all-sided leaf area on the projected ground surface [Chen et al., 1997], and it is a plant trait that is physically measurable via litter-fall.
collection, hemispheric photography and ground based optical techniques. Several global LAI products have been released using various satellite data [Myneni et al., 2002; Deng et al., 2006; Baret et al., 2013; Zhu et al., 2013] and have been incorporated into global simulations of GPP and ET [Ryu et al., 2011; Chen et al., 2012a].

In contrast to the wide use of remotely sensed LAI, the study of physiological traits mostly stays at local scales [Ustin and Gamon, 2010]. The physiological traits acquired from reflectance are generally classified into non-pigment traits [Kokaly et al., 2009] and pigment traits [Ustin et al., 2009] dependent on their sensitive bands.

Non-pigment traits include lignin, protein, cellulose and total nitrogen. The vibration frequency and length of bonds involving N, C, O, and H influence the absorption in certain bands, usually at longer wavelength (>2100 nm). However, these signals are relatively weak and are often confounded by scattering due to structural elements; therefore, most derivations of the non-pigment traits rely on the whole reflectance spectrum of vegetation [Kokaly et al., 2009; Ollinger, 2011]. Singh et al. [2015] used a partial least square regression (PLRS) statistical method to generate spatially explicit maps of seven non-pigment traits (e.g. leaf mass per area, and percentage nitrogen, lignin and cellulose) of a forest. Ollinger et al. [2008] found a strong correlation between foliage nitrogen content and a broadband canopy albedo, but it has been argued that correlation is an artifact due to the variations in canopy structure [Knyazikhin et al., 2013]. Ainsworth et al. [2014] and Serbin et al. [2015] even used the PLRS method to successfully build an empirical connection between the measured Vcmax25 and spectral characteristics. So far, the limited number of studies on the non-pigment constituents can either only hold up at local scale or lack a physical basis. It is still challenging to map non-pigment traits for large regions.

Major pigments in plants include leaf chlorophyll, carotenoid and anthocyanins. Chlorophyll a and b are responsible for harvesting light in blue (400-500 nm) and red (650-700 nm) regions, respectively. Carotenoids, including carotenes and xanthophylls, also absorb photons in the blue band to support photosynthesis. In the xanthophyll cycle, excessive solar radiation received by leaves causes a green shift in leaf reflectance.
around 531 nm [Gamon et al., 1990]. Anthocyanins are protective pigments in cytoplasm that absorb similar wavelengths as chlorophyll and protect plants from excess radiation [Ollinger, 2011]. Based on the absorbance of these pigments, considerable effort has been invested in the derivation of pigment contents from remotely sensed surface reflectance. Most studies are focused on the derivation of chlorophyll since it is the dominant pigment in plants, and the spectral overlap makes its signal difficult to disentangle from those of other pigments. Dozens of indices have been developed to represent leaf chlorophyll content (LCC) [Gitelson, 2005; Wu et al., 2008; Croft et al., 2014], using narrow red-edge bands where LCC has the strongest absorbance and the least overlap with other pigments [Gitelson et al., 1996; Knyazikhin et al., 2013]. Some studies have attempted to investigate the epoxidation state of xanthophyll cycle pigments by utilizing the green shift in reflectance [Gamon et al., 1997, 2016]. However, unlike LAI, leaf pigments are not yet widely utilized in carbon and water flux simulations by TBMs. There are two major reasons for this:

1. The complexity in the relationship between satellite-derived canopy reflectance and leaf pigments has hampered regional and global efforts to map pigments, because the vegetation structural effect confounds the identification of absorption features in foliage [Knyazikhin et al., 2013]. At the leaf level, variations in internal leaf structure, leaf thickness and water content between different plant functional types differentially affect leaf reflectance factors for a given chlorophyll content [Serrano, 2008; Croft et al., 2014]. At the canopy level, the satellite vegetation reflectance contains non-chlorophyll signals from tree architecture, leaf area index (LAI), clumping, leaf angle distribution and tree densities [Haboudane et al., 2002; Ollinger et al., 2008; Simic et al., 2011]. To remove the effects of canopy structure on LCC derivation, some studies attempted to develop physically-based models that can explicitly connect biophysical variables and canopy reflectance through modelling radiation transfer within a canopy and leaf [Croft et al., 2013; Houborg et al., 2015a]. Inversions of the well-recognized canopy radiative transfer model (i.e. SAIL, 4-SCALE) and leaf radiative model (i.e. PROSPECT) usually form the basis of such physically-based models.
2. More importantly, the pathways to incorporate pigment into the current biochemical models for photosynthesis are not clear. Even though it is known that chlorophyll captures photons to provide energy for photosynthesis, the connections between LCC and the key photosynthetic parameters such as Vcmax25 and Jmax25 have rarely been studied. Recently, Croft et al. [2017] found a single relationship between LCC and Vcmax25 for four tree species in a temperate forest. It is suggested that chlorophyll molecules are rich in nitrogen; therefore, leaf chlorophyll content (LCC) -- defined as total chlorophyll a/b per unit leaf area -- can serve as a visible indicator of the status of the photosynthetic nitrogen pool. In this way LCC can be linearly related to leaf Rubisco content and Vcmax25 [Evans, 1989; Hikosaka and Terashima, 1996]. Some studies [Houborg et al., 2013, 2015b] also developed linear relationships between Vcmax25 and LCC for C3, C4 and cropland species, using leaf nitrogen content as an intermediate.

In conclusion, this section introduces the structural and physiological traits of plants, and their roles and parameterization in TBMs. With the current progress in satellite derivation of pigments and the improved understanding of the role of pigments in photosynthesis, it is possible to incorporate satellite-derived pigments in TBMs to improve large-scale ET and GPP estimations.

1.4 Boreal Ecosystem Productivity Simulator

1.4.1 Definition of Terrestrial Biosphere models (TBMs)

The boreal ecosystem productivity simulator (BEPS) will be used in this study. BEPS falls into previously mentioned genre of models called Terrestrial Biosphere Models (TBMs). State-of-the-art TBMs are the major tools to study and simulate biogeochemical cycles (e.g. those of carbon, water, nitrogen) between the biosphere and the atmosphere, especially those regulated by biological factors (e.g. vegetation, microbes) [Moorcroft, 2006]. It is referred to as dynamic global vegetation model (DGVM) but keeps evolving by integrating more processes such as those in hydrology and soil physics and biology [Fisher et al., 2014].
TBMs are different than the land surface models (LSMs) or land surface schemes, even though they both focus on the processes on the land surface. TBMs are diagnostic and standalone models driven by observed meteorological and biophysical parameters. The fluxes simulated by TBMs in any timestep would not impact the atmospheric forcing in the subsequent time step. LSMs are prognostic models often directly coupled to atmospheric and ocean dynamics as a component of general circulation models (GCMs). The main function of LSMs is to characterize the boundary conditions for the land–atmosphere exchange fast enough to synchronize with the atmospheric dynamics. Therefore, the biophysical processes considered in LSMs are simpler than those in TBMs [Fisher et al., 2014]. Some GCMs have updated their LSMs using the principles developed in TBMs for vegetation, such as the renowned community land surface model (CLM). These upgraded GCMs are now branded as Earth system models (ESMs) if they incorporate additional components such as the oceans and ice sheets [Cox et al., 2000].

Since TBMs were originally developed to diagnose the land-atmosphere CO₂ exchange, they are generally classified into two types – enzyme kinetic (EK) models and light use efficiency (LUE) models – based on the photosynthesis module used in TBMs [Schaefer et al., 2012]. In EK models, leaf-level Rubisco kinetic and electron transport place limits on simulated GPP [Farquhar et al., 1980]. Typical EK models also have a stomatal conductance module to balance carbon uptake against water loss through the stomata, as discussed in Section 2.3. LUE models use a simple framework by estimating either GPP or NPP by multiplying LUE by absorbed photosynthetic radiation [Monteith, 1972; Field et al., 1995; Sims et al., 2006]. LUE models are adopted in many remote sensing products for their simplicity. However, the unclear role of LUE in biochemical processes and the links between the variations of LUE and climatic factors dent their credibility as a process-based model. Schaefer et al. [2012] compared 26 TBMs on their ability to simulate GPP over 39 EC towers. Only 7 out of the 26 models were LUE models and EK models generally rendered smaller root mean square error for GPP than did LUE models.

In addition, TBM are constantly evolving to include more land surface processes, such as lateral flows [Govind et al., 2009], radiation transfer inside canopies [Dai et al., 2004] and soil thermal dynamics [Ju et al., 2006]. These processes are all linked to
photosynthesis through different pathways and should be considered as an integral part of TBM s. The biochemical EK model provides a strong framework to incorporate these processes in TBM s. The TBM used in this study – the Boreal Ecosystem Productivity Simulator (BEPS) – is a state-of-the-art EK model.

1.4.2. Development of BEPS

BEPS was first developed to estimate carbon uptake [Chen et al., 1999] and ET over the Canadian landmass [Liu et al., 2003]. BEPS is characterized by a two-leaf radiation regime to consider the instantaneous radiative environment of leaves [Norman, 1982] and an analytic daily integration scheme to aggregate instantaneous estimates over a daily time step [Chen et al., 1999]. BEPS has been substantially updated since its original release, including upgrading the time step from daily to hourly [Ju et al., 2006], adding a mechanistic soil moisture and thermal scheme [Ju et al., 2006; He et al., 2014], improving the energy balance and snow simulation module [Chen et al., 2007], and designing a scheme to consider the connection between the photosynthetic capacity and the nitrogen gradient [Chen et al., 2012b]. In 2014, BEPS was intensively re-modularized to accomplish a clean and user-friendly structure. In this study, we have used the new version of hourly BEPS (version 4) to simulate ET and GPP at the site level and at the global scale.

BEPS is a recognized TBM that has participated in several international efforts to compare model performance in GPP and ET modelling [Amthor et al., 2001; Potter et al., 2001; Grant et al., 2006; Keenan et al., 2012a; Schaefer et al., 2012]. Most assessments have identified BEPS in the top quartile because it is an advanced two-leaf enzyme kinetic model. In addition, several site-level validations are conducted to evaluate BEPS performance [Chen et al., 1999; Liu et al., 2003; Ju et al., 2006; Gonsamo et al., 2013] to show that BEPS can produce GPP and ET estimates close to EC measurements. Its usage has expanded from boreal ecosystems to other plant functional types and it thereby produces reliable global estimates [Wang et al., 2004; Chen et al., 2012b; Keenan et al., 2012b]. Therefore, BEPS provides a suitable platform to incorporate new physiological and structural traits and enables exploration of their influence on global GPP and ET simulations.
1.5 Roadmap of the thesis

An accurate understanding of the earth system is imperative to understanding implications of increasing greenhouse gas concentrations on regional and global climate. Terrestrial evapotranspiration (ET) is a major flux between the biosphere and the atmosphere. It plays a critical role in the water balance and energy balance of the land surface. In the past decade, the scientific community has made substantial progress in determining the global terrestrial ET; however, we still face challenges in understanding and simulating the variability and trend of ET and its components.

This study will focus on improving the estimation of transpiration, which accounts for 55-80% of terrestrial ET. Water and carbon exchanges between vegetation and the atmosphere are known to be coupled through leaf stomata; therefore, an improved understanding of transpiration will also lead to improvement in photosynthesis estimation. Transpiration is controlled by a diverse range of biotic and abiotic factors, which are linked to structural and physiological status of vegetation. Current terrestrial biosphere models usually misrepresent key vegetation traits that influence carbon and water fluxes (e.g. Vcmax25) due to the difficulty of quantifying them over large scales and the lack of knowledge about the function of these traits in ecosystems. This study strives to solve this misrepresentation of vegetation status in TBMs and in turn improves the simulation of terrestrial ET. This study is comprised of three parts:

In Chapter 2, the key structural parameters leaf area index (LAI) and clumping index (CI) are used to improve a two-leaf upscaling scheme to replace the commonly used big-leaf scheme and two-big-leaf scheme in TBMs. The two-leaf scheme was initially developed and evaluated for GPP simulation but was not comprehensively tested for ET simulation. The two-leaf upscaling scheme has the advantage of simulating the instantaneous leaf environment and avoiding the uncertainties propagated from artificially calculating canopy conductance (g_c). This study is the first of this kind to demonstrate the advantage of the two-leaf scheme on ET modelling over a spectrum of sites.
In Chapter 3, a long-recognized physiological parameter – leaf chlorophyll content (LCC) is added to the TBM to represent the varying photosynthetic capacity over the season and space, based on our previous finding that that LCC is linearly related to Vcm25 of leaves. This chapter also updates the two-leaf scheme accordingly for the purpose of including LCC. The updated two-leaf scheme thus enables the analysis of the physiological difference between sun-adapted and shade-adapted leaves. This study is the first attempt to use LCC in a TBM to simulate GPP and ET at a natural site.

Following the conclusion drawn from site-level studies in Chapters 2 and 3, Chapter 4 will use remote sensing techniques to derive LCC, LAI and CI for flux sites over the globe. Then we use these parameters in BEPS and analyse how they, especially LCC, will improve ET and GPP simulations for different plant functional types. This lays the foundation for the future work of producing more accurate global GPP and ET maps.

Chapter 5 summarizes the main results of this thesis, discusses the uncertainty involved in using satellite derived vegetation status in modelling fluxes and reflects on directions for future improvements.

1.6 References


Jiang, C., and Y. Ryu (2016), Multi-scale evaluation of global gross primary productivity


soil and vegetation energy fluxes in observations of directional radiometric surface
temperature, *Agricultural and Forest Meteorology*, 77(3), 263–293, 
doi:https://doi.org/10.1016/0168-1923(95)02265-Y.

Oki, T., and S. Kanae (2006), Global hydrological cycles and world water resources.,

Ollinger, S. V et al. (2008), Canopy nitrogen, carbon assimilation, and albedo in
temperate and boreal forests: Functional relations and potential climate feedbacks,
*Proceedings of the National Academy of Sciences*, 105(49), 19336–19341.

Ollinger, S. V. (2011), Sources of variability in canopy reflectance and the convergent
8137.2010.03536.x.

Elsevier.

Papale, D. et al. (2006), Towards a standardized processing of Net Ecosystem Exchange
measured with eddy covariance technique: algorithms and uncertainty estimation,

Moreno, C. Frankenberg, and J. A. Berry (2014), Linking chlorophyll a
fluorescence to photosynthesis for remote sensing applications: Mechanisms and
challenges, *Journal of Experimental Botany*, 65(15), 4065–4095,

Potter, C. S. et al. (2001), Comparison of boreal ecosystem model sensitivity to
variability in climate and forest site parameters, *Journal of Geophysical Research*,

Reichstein, M. et al. (2005), On the separation of net ecosystem exchange into
assimilation and ecosystem respiration: Review and improved algorithm, *Global

Richardson, A. D. et al. (2006), A multi-site analysis of random error in tower-based
measurements of carbon and energy fluxes, *Agricultural and Forest Meteorology*,

Rogers, A. (2014), The use and misuse of Vc,max in Earth System Models,

Netherlands, Dordrecht.

Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto
(2004), A Continuous Satellite-Derived Measure of Global Terrestrial Primary
Production, *BioScience*, 54(6), 547, doi:10.1641/0006-

Ryu, Y. et al. (2011), Integration of MODIS land and atmosphere products with a
coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales, *Global Biogeochemical Cycles*, 25(4), n/a-n/a, doi:10.1029/2011GB004053.


Tang, X. et al. (2015), How is water-use efficiency of terrestrial ecosystems distributed and changing on Earth?, *Scientific Reports, 4*(1), 7483, doi:10.1038/srep07483.


INTRODUCTION.


Chapter 2

Comparison of big-leaf, two-big-leaf and two-leaf upscaling schemes for evapotranspiration estimation using coupled carbon-water modelling

Preamble: this chapter is under review by *Journal of Geophysical Research – Biogeoscience*. The authors of this paper are: X. Luo, J. M. Chen, J. Liu, T. A. Black, H. Croft, R. Staebler, L. He, M. A. Arain, B. Chen, G. Mo, A. Gonsamo, H. McCaughey. XL performed the data analysis and wrote the paper, JC and JL developed the idea and helped with the writing. TB, HC, RS, MA, AG and HM processed and provided the flux data and ancillary data. XL, LH, BC and GM updated the code of BEPS model.

2.1 Abstract

Evapotranspiration (ET) from plant canopies is commonly estimated using the Penman-Monteith (PM) equation, which assumes that the plant canopy is a big leaf. One critical issue of applying the big-leaf scheme (*BL*) is to upscale canopy stomatal conductance (*Gₚ*) from leaf stomatal conductance (*gₛ*). However, *BL* has been found to be unsuitable for terrestrial biosphere models (TBMs) built on the carbon-water coupling principle because it fails to capture daily variations of gross primary productivity (GPP). A two-big-leaf scheme (*TBL*) and a two-leaf scheme (*TL*) that stratify a canopy into sunlit and shaded leaves have been developed to address this issue for GPP. However, there is a lack of comprehensive comparison of these upscaling schemes for ET estimation, especially on the difference between *TBL* and *TL*. This study uses eddy-covariance measurements from nine sites in Canada to compare the performance of *BL*, *TBL* and *TL*.

We find that *TL* shows strong performance (*r²=0.71, slope=0.91, RMSE=0.05 mm/h) in estimating ET across nine eddy-covariance towers. *BL* underestimates annual GPP and ET, and *TBL* tends to underestimate ET and overestimate GPP. The biases of estimates from *BL* and *TBL* increases with leaf area index (LAI) but the biases from *TL* show no trends with LAI. *BL* miscalculates the separation of light-saturated and light-unsaturated leaves in canopies, incurring errors in GPP estimates and the consequent underestimation of ET. *TBL* mainly underestimates sunlit GPP and shaded ET and overestimates sunlit
ET and shaded GPP. The relative underestimation of fluxes in TBL to TL is driven by the questionable calculation of $G_s$ in TBM$^s$ while the overestimation is driven by the underestimation of intercellular CO$_2$ concentration in TBL. The biases of GPP and ET estimates from TBL systematically increase with LAI due to the structural errors above.

2.2 Introduction

Land surface evapotranspiration (ET) plays a critical role in the water and energy exchanges between the biosphere and the atmosphere. It accounts for 60% of the terrestrial precipitation [Oki and Kanae, 2006] and consumes 50% of the solar energy absorbed by the land surface [Trenberth et al., 2009]. In the past decades, the Penman-Monteith (PM) equation has provided a sound foundation for estimating ET from the site to the global scales [Sellers, 1986; Dickinson et al., 1993; Bonan, 1996; Moran et al., 1996; Weiß and Menzel, 2008; Mu et al., 2011; Wang and Dickinson, 2012].

The PM equation perfectly combines physical and physiological constraints into one simple equation for ET estimations [Monteith and Unsworth, 2013]. However, the simplicity of the PM equation also leads to a potential imperfection: in order to calculate canopy conductance ($G_c$), the PM equation has to use a big leaf assumption, which abstracts the whole canopy into a one-layer source. This assumption is apparently in conflict with the complex structures of canopies in reality, where the leaf distribution varies by clumping [Chen et al., 1997], light environments [Norman et al., 1982] and angle and heights [Baldocchi and Meyers, 1988] and consequently influence the canopy transpiration rates.

However, a considerable number of studies have used $G_c$ to produce reliable ET results regardless of the potential defect of the PM equation, and hence corroborated the validity of the big-leaf scheme (BL) underlying the PM equation [Dickinson et al., 1991; Moran et al., 1996; Mu et al., 2011; Yan et al., 2012; Monteith and Unsworth, 2013]. These studies regarded ET as an independent process, and $G_c$ for the PM equation can be freely tuned to fit the ET measurements. $G_c$ is acquired through either a top-down or a bottom-up method in ET studies. In the top-down method, $G_c$ is derived by inverting the PM
equation using near-surface measurements of the latent heat flux and meteorological variables [Stewart, 1988; Kelliher et al., 1995; Phillips and Oren, 1998; Lai et al., 2000; Monteith and Unsworth, 2013]. The reciprocal of $G_c$ value represents the bulk resistance enforced collectively by leaf stomata and soil to transport water [Raupach and Finnigan, 1988; Paw U and Meyers, 1989]. Process models used for large-scale simulations of ET are often equipped with the bottom-up method, which identifies “two layers” for ET, namely the transpiration from vegetation and the evaporation from soil. An integrated canopy stomatal conductance ($G_s$) is used to represent the control of vegetation in such two-layers models [Norman et al., 1995]. Several theoretical and experimental studies have suggested that $G_s$ is not equivalent to $G_c$, though the value of $G_s$ would be close to $G_c$ for dense vegetation [Kelliher et al., 1995; Baldocchi and Meyers, 1998] and $G_s$ is also widely used in the PM equation to calculate transpiration.

However, with the emergence of process-based Terrestrial Biosphere models (TBMs) that consider carbon and water exchange as a coupled process, $G_s$ acquired from BL should be able to satisfy the simulation of ET as well as simulation of the carbon uptake simultaneously. The statistical model or semi-empirical models that quantify $G_c$ or $G_s$ by inversing ET measurements or using empirical indices would no longer suffice for TBMs. The concept of $G_c$ and $G_s$ may not be useful anymore because photosynthesis model is only developed for leaves not for canopies. Ball et al.[1987] and Leuning [1990] discovered that stomatal conductance ($g_s$) is linearly tuned by the carbon assimilation rate ($A$) of leaves, denoted “Ball-Woodrow-Berry stomatal conductance model” here. Sellers et al. [1992] and Amthor [1994] made the first efforts to update BL for TBMs. They assumed $A$ decreases from the top to the bottom of a canopy following either the foliage nitrogen gradient or long-term solar radiation gradient, and so does $g_s$. These gradients are expressed in a form of an exponential function dependent on the canopy depth which is quantified using the accumulated LAI from the canopy top. Afterward, the canopy total photosynthesis ($A_c$, a.k.a. GPP) can be easily upscaled from $A$ using these functions and then $G_s$ is calculated through the Ball-Woodrow-Berry stomatal conductance model.
BL designed for the carbon-water coupled TBMs were shown to perform well at some sites, but many researchers reported an underestimation of GPP by these models, since $A$ is more sensitive to the instantaneous solar radiation on leaves, while nitrogen and the long-term radiation gradient cannot explain the rapid changes in $A$ as described in BL [De Pury and Farquhar, 1997; Friend, 2001]. For example, a leaf at the bottom of a canopy in a sun fleck will instantaneously receive far more radiation for photosynthesis than the average radiation that Beer’s law would predict. To describe the instantaneous radiation intercepted by leaves, a two-leaf radiation regime was developed [Sinclair et al., 1976; Norman, 1982; De Pury and Farquhar, 1997; Chen et al., 1999]. It separates a canopy into a group of sunlit leaves and a group of shaded leaves. $A$ of a sunlit leaf tends to be light-saturated by receiving both direct and diffuse solar radiation, while $A$ of a shaded leaf is capped by the amount of diffuse radiation on leaves. Based on the two-leaf radiation regime, a hierarchy of upscaling schemes including the multi-layer scheme, the two-big-leaf scheme (TBL) and the two-leaf scheme (TL) are developed for TBMs.

Leuning et al. [1995] and Baldocchi and Harley [1995] developed the multi-layer scheme, in which a canopy is separated into layers, and every layer is divided into sunlit and shaded segments. The multi-layer scheme considers the ecological processes inside the canopy in great detail: leaf nitrogen, leaf photosynthetic capacity and even leaf inclination angles can be prescribed independently. In this scheme, the leaf photosynthesis and transpiration are calculated for each segment and then integrated into the canopy-scale GPP and ET by multiplying by the LAI of each segment. Though the multi-layer scheme is regarded as the most accurate way to upscale fluxes from leaf to canopy, its expensive computational demand for large-scale applications drives the need to use simple upscaling schemes in TBMs [Wang and Leuning, 1998].

Some studies then developed an upscaling scheme which is classified as TBL, inheriting the idea of BL and using the two-leaf radiation regime [De Pury and Farquhar 1997; Wang and Leuning, 1998; Dai et al., 2004; Ryu et al., 2011]. $A_c$ and $G_s$ for sunlit and shaded canopies are simulated respectively in TBL, and $G_s$ of each leaf group is then used in the PM equation to calculate ET. In order to calculate $A_c$ and $G_s$, TBL requires the biochemical parameters of leaves to be upscaled to their canopy counterparts. Since the biochemical
model (i.e. Farquhar’s biochemical model) is originally developed to simulate leaf-level photosynthesis, the direct application of it at the canopy scale may bring unexpected uncertainties in simulation when the physiological behaviour of an imaginary “big leaf” can surpass the explanatory ability of a leaf-level model.

To avoid the uncertainties involved in the artificial upscaling of leaf parameters introduced in TBL, there is a need to circumvent the use of $G_s$. One apparent solution is to directly use $g_s$ in the PM equation and simulate leaf transpiration and upscale it to canopy-scale ET by multiplying by LAI, as it is shown in the multi-layer scheme. Chen et al. [1999, 2012] further developed TL as an alternative to the expensive multi-layer scheme and TBL. They found that the most distinguishable environmental factor in leaf micrometeorology is the intercepted solar radiation; therefore, we only need to separate the canopy into two segments (i.e. sunlit and shaded) and calculate the $A$ and $g_s$ of a representative leaf from each segment. This method takes advantage of the two-leaf radiation regime and avoids the use of $G_s$ in TBMs. It is also conceptually rigorous in running Farquhar’s biochemical model, the Ball-Woodrow-Berry stomatal conductance model and the PM equation simultaneously at the leaf level, since the first two were developed based on leaf-level measurements.

Since the application of the two-leaf radiation regime in TBMs in 1990s, some studies have strived to evaluate the performance of different upscaling schemes with flux measurements. The advantage of TBL over BL has been proved at two flux sites for GPP modelling [Medlyn et al., 2003; Mercado et al., 2006], and TL has been validated with data from 11 eddy-covariance (EC) towers and proved its advantage over BL for GPP modelling [Sprintsin et al., 2012]. As for ET modelling, Vogel et al. [1995] has used a TBM with the multi-layer scheme to simulate ET and compared it with a hierarchy of less-sophisticated ET models over a well-irrigated cropland, and suggested no advantage of using the complicated multi-layer scheme for ET modelling. This conclusion may not be applicable for TBMs since the parameters for those less sophisticated ET models can be arbitrarily tuned to fit the measurements whereas the parameters of TBMs are simulated based on the physiological principle of carbon-water coupling. Currently, we still lack a clear understanding of the effects of upscaling schemes in TBMs for ET simulations and
how these effects vary across sites. In addition, there is a need to clarify the definitions of the two-leaf radiation regime, TBL and TL because of their interchangeable uses in previous studies [De Pury and Farquhar, 1997; Wang and Leuning, 1998]. Therefore, the objective of this research is to compare BL, TBL and TL over a spectrum of flux sites and analyse their influence on ET modelling.

2.3 Data and method

2.3.1 Description of the model

BEPS is an enzyme kinetic, two-layers (i.e. vegetation and soil) and dual-source (sunlit and shaded) model first developed to estimate carbon uptake and the water cycle over the Canadian landmass [Liu et al., 2003]. It is characterized by a two-leaf separation strategy [Norman, 1982] and an analytic daily integration scheme [Chen et al., 1999]. Several inter-model comparisons and site-level validations have shown that BEPS can produce reasonable GPP and ET estimates [Amthor et al., 2001; Potter et al., 2001; Liu et al., 2003; Grant et al., 2006]. Its usage has expanded from boreal ecosystems to other plant functional types in the past decade [Wang et al., 2004; Chen et al., 2012; Gonsamo et al., 2013] and BEPS has been updated to support simulations at hourly and half-hourly steps [Chen et al., 2007].

In BEPS, ET from the land surface mainly consists of three components: transpiration from leaves, evaporation (sublimation) from the wet canopy and evaporation (sublimation) from the soil surface. Since this study focuses on leaf-to-canopy upscaling methodologies and their effects on ET estimation, we will primarily describe the transpiration-related processes in BEPS.

According to TL, BEPS simulates the photosynthetic rate of a representative sunlit leaf \( A_{sunlit} \) and a shaded leaf \( A_{shaded} \) first, and then obtains the canopy photosynthetic productivity \( A_c \) as the sum of the photosynthesis from sunlit leaves and shaded leaves (Equation 1). Similar to the upscaling of \( A_c \), the transpiration of the canopy \( T_c \) is the sum of transpiration from these two groups of leaves (Equation 2). The two-leaf scheme
assumes that all sunlit leaves (shaded leaves) are exposed to the same environment (i.e. irradiance, temperature, vapor pressure deficit) and have the same physiological features (i.e. $V_{\text{max}}^{25}$), and therefore the transpiration and photosynthesis of the whole leaf group can be predicted using one representative leaf.

$$A_c = A_{\text{sunlit}} \times LAI_{\text{sunlit}} + A_{\text{shaded}} \times LAI_{\text{shaded}} \quad (2-1)$$

$$T_c = T_{\text{sunlit}} \times LAI_{\text{sunlit}} + T_{\text{shaded}} \times LAI_{\text{shaded}} \quad (2-2)$$

where $A_{\text{sunlit}}$ and $A_{\text{shaded}}$ are the photosynthetic rates of a representative sunlit leaf and a representative shaded leaf, respectively. They are acquired from an analytic solution derived from a leaf biochemical model and a mass transfer equation for the leaf [Baldocchi, 1994]. The maximum carboxylation velocity ($V_{\text{max}}^{25}$) and the maximum electron transport capacity ($J_{\text{max}}^{25}$) at 25 °C for sunlit and shaded leaves are calculated based on a nitrogen gradient in the canopy (Appendix 2.7.1) to parameterize the biochemical processes in BEPS. $T_{\text{sunlit}}$ and $T_{\text{shaded}}$ are the transpiration from sunlit leaf and shaded leaf, respectively. $LAI_{\text{sunlit}}$ and $LAI_{\text{shaded}}$ are the LAI of sunlit leaves and shaded leaves, respectively. The values of $LAI_{\text{sunlit}}$ and $LAI_{\text{shaded}}$ are calculated following the stratification scheme of Norman [1982] and Chen et al. [1999].

$$LAI_{\text{sunlit}} = 2 \cos \theta (1 - \exp(-0.5\Omega LAI_{\text{tot}} / \cos \theta)) \quad (2-3)$$

$$LAI_{\text{shaded}} = LAI_{\text{tot}} - LAI_{\text{sunlit}} \quad (2-4)$$

where $\theta$ is the solar zenith angle, $LAI_{\text{tot}}$ is the total leaf area index of the canopy and $\Omega$ is the clumping index.

Then, the PM equation is employed to calculate $T$ of a sunlit or shaded leaf (Equation 5)

$$T = \frac{\Delta(R_n - G) + \rho c_p VPD g_v}{\Delta(1 + \frac{\Delta V}{g_s}) \gamma} \times \frac{1}{\lambda} \quad (2-5)$$

where $\lambda$ is the latent heat of evaporation of water, $R_n$ is the net radiation at the leaf surface (Appendix 2.7.2), $G$ is the heat storage of the leaf which can be neglected, $\rho$ is the density of air, $c_p$ is the specific heat of air, $VPD$ is the vapor pressure deficit of the
ambient air, γ is the psychrometric constant, $g_{w}$ is leaf boundary layer conductance for water vapor, $\Delta$ is the slope of the saturation vapor pressure curve at air temperature, and $g_s$ is the stomatal conductance of the representative sunlit or shaded leaf.

A modified Ball-Woodrow-Berry model is then used to calculate the $g_s$ of sunlit or shaded leaves [Chen et al., 2012], respectively.

$$g_s = f_w \left( \frac{m A}{C_s} \right) + g_0$$  \hspace{1cm} (2-6)

where $m$ is the dimensionless Ball-Woodrow-Berry coefficient, set at 8 for C3 plants, $RH$ is the relative humidity, $C_s$ is the carbon dioxide concentration on the leaf surface, $g_0$ is the minimum conductance at night, and $A$ is the rate of photosynthesis ($\mu$mol/m$^2$/s) of the representative sunlit or shaded leaf. The variable $f_w$, which is the soil water stress factor, is added to overcome the inability of the Ball-Woodrow-Berry equation to close the stomata during drought spells. It is widely employed as a complementary parameter to represent the regulation of the conductance of water through stomata [Sala and Tenhunen, 1996; Xu and Baldocchi, 2003]. BEPS has developed a mechanistic module to simulate soil moisture and $f_w$ [Ju et al., 2006]. However, sometimes the performance of the soil moisture module is biased because the module requires high accuracy of the parameterization of soil texture in multiple layers. To minimize the possible deviations in $g_s$ caused by the soil moisture simulation, we replaced the soil moisture module with measured soil moisture in this study and applied a simple equation to calculate $f_w$ (Appendix 2.7.3). The incorporation of measured surface soil moisture also reduces the errors in the estimates of surface evaporation. With this modification, the overall change in ET between schemes is mainly attributed to the transpiration, and in turn can be attributed to the corresponding upscaling scheme.

2.3.2 Modelling schemes

2.3.2.1 Big-leaf scheme (BL)
BL developed by Sellers et al. (1992, 1997) is one of the first attempts to simulate water and carbon fluxes simultaneously, in which

\[
A_c = A_0 \times \frac{1 - \exp(-k \times LAI_{tot})}{k}
\]  

(2-7)

where \(A_0\) is the photosynthetic rate of the leaves at the top of the canopy, and \(A_c\) is the total canopy photosynthesis rate. Since BL assumes an optimal nitrogen gradient following the long-term solar radiation gradient, \(k\) is the extinction coefficient for both solar radiation and nitrogen gradients in a canopy and it is set at 0.5. After obtaining \(A_c\), \(g_c\) for the big-leaf is then acquired using the Ball-Woodrow-Berry equation introduced in Equation (6). To facilitate our analysis, \(g_c\) is simplified into the form of:

\[
G_s = g_{s0} \times \frac{1 - \exp(-k \times LAI_{tot})}{k}
\]  

(2-8)

where \(g_{s0}\) is the stomatal conductance of the leaves on top of the canopy.

2.3.2.2 Two-big-leaf scheme (TBL)

TBL applies a different way of describe the dual sources than TL. The TBL scheme requires an artificial upscaling of leaf-level physiological parameters \(V_{\text{max}}^{25}\) and \(J_{\text{max}}^{25}\) to their counterparts for each leaf group (i.e. \(V_{\text{max,sunlit can}}^{25}\), \(V_{\text{max,shaded can}}^{25}\), \(J_{\text{max,sunlit can}}^{25}\), \(J_{\text{max,shaded can}}^{25}\)). By incorporating these canopy-scale parameters into Farquhar’s model and the Ball-Woodrow-Berry equation we are able to obtain \(A_c\) and \(g_c\) for the sunlit and shaded leaf groups, respectively. For the purpose of this study, TBL is added to existing BEPS to compute the \(G_s\) of sunlit leaves (\(G_{s,sunlit}\)) and shaded leaves (\(G_{s,shaded}\)) [Wang and Leuning, 1998; Dai et al., 2004; Ryu et al., 2011]. The calculation of the canopy-scale \(V_{\text{max}}^{25}\) in TBL is introduced in Appendix 2.7.4.

2.3.2.3 Two-leaf scheme (TL)

The default BEPS uses TL to estimate the transpiration of a representative sunlit leaf and a representative shaded leaf first, and then upcales the leaf-level transpiration to canopy-level by multiplying by the corresponding LAI values (Equation 2) (see section
“Description of the model” for a more detailed description of TL). This method avoids the use of \( G_s \) and canopy-level photosynthetic parameters, so it is described as TL.

Figure 2-1. Schematic descriptions of the three upscaling schemes: BL, TBL and TL. Brightness of the colour represents the intensity of solar radiation intercepted by leaves; saturation of the colour represents the variation in N content of leaves.

2.3.3 Validation sites and input data

The data used to drive the model are obtained from Fluxnet (http://fluxnet.ornl.gov/). Nine sites in Canada are selected mainly because they have some measured leaf area index (LAI), clumping index (\( \Omega \)) and soil moisture data (Table 2-1). Using these measurements can effectively constrain the uncertainty for ET simulation than using inferred values from prognostic sub-models. The input meteorology data include incident solar irradiance (W/m\(^2\)), air temperature (°C), precipitation (mm/h), relative humidity (%), wind speed (m/s) and soil water content (m\(^3\)/m\(^3\)). Overstory LAI (\( LAI_o \)) data were measured during some growing seasons at these sites. We use the reflectance data of the Moderate-resolution Imaging Spectroradiometer (MODIS) to extrapolate the LAI measurements to daily \( LAI_o \) sequences [Gonsamo and Chen, 2014]. Except for an old aspen site (CaOas), the understorey LAI (\( LAI_u \)) is calculated using an empirical
equation, $LAI_u = 1.18 \times e^{(-0.99 \times LAI_0)}$ [Liu et al., 2003]. Since CaOas has an $LAI_u$ comparable to $LAI_o$ [Barr et al., 2004], its $LAI_u$ is calculated as 90% of $LAI_o$. $\Omega$ is also a critical canopy structural parameter, as it defines the non-randomness of the foliage distribution in a canopy (i.e. the overlapping of the leaves and aggregation of the needles in a shoot) [He et al., 2012]. $\Omega$ ranges from 0 to 1, with a higher number indicating that the canopy is closer to a random distribution.

$V_{\text{max}}^{25}$ is a critical parameter in Farquhar’s photosynthesis model. $V_{\text{max}}^{25}$ for each site is obtained from previous data assimilation work [Groenendijk et al., 2011; He et al., 2014]. The temporal variation in $V_{\text{max}}^{25}$ is also considered by assuming that the seasonal patterns of $V_{\text{max}}^{25}$ follows the season pattern of LAI [Ryu et al., 2011]. In this study, the $V_{\text{max}}^{25}$ value on a given day in a growing season is calculated using an empirical equation:

$$V_{\text{max, day}}^{25} = \alpha \times V_{\text{max}}^{25} + (1 - \alpha) \times V_{\text{max}}^{25} \times \frac{L_c - L_{\text{min}}}{\beta \times L_{\text{max}} - L_{\text{min}}}$$  \hspace{1cm} (2-9)$$

where $L_{\text{max}}, L_{\text{min}}$ and $L_c$ are maximum, minimum and current LAI values over the year, respectively. The empirical variables $\alpha$ and $\beta$ are set as 0.30 and 0.75, respectively. The ratio term $\frac{L_c - L_{\text{min}}}{\beta \times L_{\text{max}} - L_{\text{min}}}$ should range between 0 and 1.
<table>
<thead>
<tr>
<th>Site code</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Year</th>
<th>Land cover</th>
<th>Overstorey main genera</th>
<th>Maximum overstorey LAI</th>
<th>Clumping index</th>
<th>$V_{\text{max}}$ (µmol m$^{-2}$ s$^{-1}$)</th>
<th>Number of LAI measurements</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaCa2</td>
<td>49.8705</td>
<td>-125.2909</td>
<td>2002-2010</td>
<td>ENF</td>
<td><em>Pseudotsuga menziesii</em></td>
<td>4.45</td>
<td>0.48$^d$</td>
<td>38.8</td>
<td>26</td>
<td><em>Chen et al.</em> [2009]</td>
</tr>
<tr>
<td>CaCa3</td>
<td>49.5346</td>
<td>-124.9004</td>
<td>2002-2010</td>
<td>ENF</td>
<td><em>Pseudotsuga menziesii</em></td>
<td>8.15</td>
<td>0.532</td>
<td>38.8</td>
<td>31</td>
<td><em>Chen et al.</em> [2009]</td>
</tr>
<tr>
<td>CaCbo</td>
<td>44.3185</td>
<td>-79.9342</td>
<td>2008-2013</td>
<td>DBF</td>
<td><em>Acer rubrum, Populus tremuloides</em></td>
<td>4.96</td>
<td>0.72$^e$</td>
<td>62</td>
<td>30$^b$</td>
<td><em>Froelich et al.</em> [2015]</td>
</tr>
<tr>
<td>CaGro</td>
<td>48.2173</td>
<td>-82.1555</td>
<td>2005-2011</td>
<td>MF</td>
<td><em>Picea mariana</em></td>
<td>3.87</td>
<td>0.821</td>
<td>40</td>
<td>6</td>
<td><em>Gökkaya et al.</em> [2013]</td>
</tr>
<tr>
<td>CaOas</td>
<td>53.6289</td>
<td>-106.1978</td>
<td>2002-2010</td>
<td>DBF</td>
<td><em>Populus tremuloides</em></td>
<td>2.43</td>
<td>0.87</td>
<td>62.5$^f$</td>
<td>9</td>
<td><em>Barr et al.</em> [2004]</td>
</tr>
<tr>
<td>CaOjs</td>
<td>53.9872</td>
<td>-105.1178</td>
<td>2002-2010</td>
<td>ENF</td>
<td><em>Picea mariana</em></td>
<td>3.45</td>
<td>0.662</td>
<td>39.4</td>
<td>6</td>
<td><em>Bergeron et al.</em> [2007]</td>
</tr>
<tr>
<td>CaOjp</td>
<td>53.9163</td>
<td>-104.6920</td>
<td>2005-2010</td>
<td>ENF</td>
<td><em>Pinus banksiana</em></td>
<td>2.01</td>
<td>0.599</td>
<td>31</td>
<td>12</td>
<td><em>Barr et al.</em> [2006]</td>
</tr>
<tr>
<td>CaTp3</td>
<td>42.7068</td>
<td>-80.3483</td>
<td>2009-2013</td>
<td>ENF</td>
<td><em>Pinus strobus</em></td>
<td>7.17</td>
<td>0.518</td>
<td>31</td>
<td>15</td>
<td><em>Peichl et al.</em> [2010]</td>
</tr>
<tr>
<td>CaTp4</td>
<td>42.7098</td>
<td>-80.3574</td>
<td>2009-2013</td>
<td>ENF</td>
<td><em>Pinus strobus</em></td>
<td>8.10</td>
<td>0.513</td>
<td>31</td>
<td>15</td>
<td><em>Arain and Restrepo-Coupe</em> [2005]</td>
</tr>
</tbody>
</table>

a. The land cover type is adopted from the site introductions on the Fluxnet. The selected sites include deciduous broadleaf forests (DBF), evergreen needleleaf forests (ENF) and mixed forests (MF).

b. The maximum value of the available LAI measurements on the Fluxnet. These values refer total LAI of the overstorey.
c. [Chen et al., 2006]

d. Canada Carbon Program

e. [He et al., 2012]

f. $V_{max}^{25}$ values refer to the maximum carboxylation capacity at 25°C for leaves on top of canopies, and they are derived based on the values provided by Groenendijk et al (2011). We assume that same species should have similar $V_{max}^{25}$ values. For all the Douglas fir ($Pseudotsuga menziesii$) sites in British Columbia, $V_{max}^{25}$ ranges from 20.5 to 54.1 µmol m$^{-2}$ s$^{-1}$ [Groenendijk et al., 2011]. In this study, the median $V_{max}^{25}$ (38.8 µmol m$^{-2}$ s$^{-1}$) of all these sites is assigned to site CaCa2 and CaCa3; CaCbo uses the average $V_{max}^{25}$ of all temperate deciduous forest; CaGro uses the average $V_{max}^{25}$ of all temperate mixed forest; CaObs uses the value that is provided. Since the dominate species of CaGro and CaObs are black spruce ($Picea mariana$), their $V_{max}^{25}$ should be similar. CaTp4 is the only pine ($Pinus$) site with a known $V_{max}^{25}$ of 31 µmol m$^{-2}$ s$^{-1}$. Since CaTp3 and CaOjp are also pine sites, they are also assigned a $V_{max}$ value of 31 µmol m$^{-2}$ s$^{-1}$.

g. [He et al., 2014]

h. [Croft et al., 2015]
2.4 Results

2.4.1 Annual ET and GPP estimates from each scheme

Three versions of BEPS using different upscaling schemes (BL, TBL, and TL) are used to simulate ET and GPP at nine eddy-covariance (EC) sites, and results from each scheme are evaluated against tower measurements (Figure 2-2; Appendix 2.7.5).

Figure 2-2 compares the estimated annual ET from each scheme with the EC measurements. Averaging across sites, annual ET are 286 mm yr\(^{-1}\), 318 mm yr\(^{-1}\), 340 mm yr\(^{-1}\), and 325 mm yr\(^{-1}\) for BL, TBL, TL and EC measurements, respectively, whereas the observational uncertainty of ET measurements is 2.5 mm yr\(^{-1}\). The annual ET estimated by BL and TBL are 16% and 7% lower compared to TL. BL, TBL and TL produce similar evaporative fluxes from soil, indicating that soil evaporation is largely determined by the total radiation incident on the ground. Most of the difference between TL and TBL is caused by shaded leaves, where the average difference in ET estimates between TL and TBL is 24 mm yr\(^{-1}\), while the difference between sunlit ET estimates of TBL and those of TL is only about -2 mm yr\(^{-1}\).
**Figure 2-2.** Simulated and measured annual ET at the studied sites, as well as the ET components obtained using \( BL, \ TBL \) and \( TL \).

Figure 2-3 demonstrates that GPP is underestimated by \( BL \) at all sites, while the GPP estimates from \( TBL \) and \( TL \) show a complex relationship. Five out of the nine sites have smaller GPP estimates from \( TBL \) than those from \( TL \), while four sites show the opposite. The pattern is clearer when partitioning GPP into its sunlit and shaded components: at eight out of the nine sites, \( TL \) produces higher sunlit GPP than \( TBL \); at seven sites, shaded GPP from \( TL \) is smaller than that from \( TBL \). Averaged across all sites, the total GPP are 922 g C m\(^{-2}\) y\(^{-1}\), 1250 g C m\(^{-2}\) y\(^{-1}\), 1232 g C m\(^{-2}\) y\(^{-1}\), and 1165 g C m\(^{-2}\) y\(^{-1}\) for \( BL, \ TBL, \ TL \) and EC measurements, respectively, whereas the uncertainty of annual GPP is around 6.3 g C m\(^{-2}\) y\(^{-1}\). Compared to \( TL \), \( BL \) underestimates annual GPP by 25% and \( TBL \) overestimates GPP slightly by 1.5%.

**Figure 2-3.** Simulated and measured annual GPP at the study sites, as well as the GPP components obtained using \( BL, \ TBL \) and \( TL \).
2.4.2 Biases of ET and GPP estimates

*BL* has been used in several carbon-water coupled TBMs [Cramer et al., 2001; Alton et al., 2007]. Some studies have noticed the underestimation of GPP by *BL*, but the accompanying underestimation of ET has been less reported and the reason for its underestimation was inadequately studied across sites. The differences between the ET and GPP estimates of *TL* and *TBL* have not been studied as well. Figure 2-4 demonstrates the biases of annual ET and GPP estimates from *BL*, *TBL* and *TL* and their relationships with LAI.

![Graphs showing biases of annual ET and GPP estimates](image)

**Figure 2-4.** The biases of estimated annual ET (a) and GPP (b), and the relative biases (biases normalized by measured fluxes) of estimated annual ET (c) and GPP (d) along with site mean LAI for each upscaling scheme.

According to Figure 2-4, *BL* underestimates both GPP and ET, and *TBL* tends to underestimate ET but overestimate GPP. The biases of estimates from *BL* and *TBL*...
change significantly (p<0.05) with LAI, indicating that sites with dense foliage are inclined to create large errors in ET and GPP estimates for BL and TBL. In contrast, the biases of ET and GPP estimates are small and insensitive to LAI. For those low-LAI sites, the difference between TBL and TL are negligible, but their difference enlarges for high-LAI sites. The results suggest the existence of errors in the structural of BL and TBL which can drive biases to systematically increase with LAI. In addition, the errors incurred by TBL are smaller than those incurred by BL for ET and GPP estimates.

2.4.3 Difference between the radiation regimes in BL and TBL(TL)

BL uses Beer’s law to describe the radiation distribution inside a canopy, while TBL and TL both use the two-leaf radiation regime to describe the radiation distribution. The quantification of the intercepted radiation of leaves affects the photosynthetic rates of leaves, and consequently influences conductance and ET. Figure 2-5 demonstrates the amount of light-saturated leaves for a given sunlit day at the nine sites.
Figure 2-5. The amount of light-saturated leaves (LAI) at each site using BL and TBL (TL). The light saturation point is fixed at 400 W/m² for this analysis. Coordinated Universal Time (UTC) is used for abscissas.

Figure 2-5 shows that BL usually classifies more leaves as light-saturated leaves than TBL and TL. Considering that light-saturated leaves have high photosynthetic rates, the GPP and ET contributed from light-saturated leaves are larger in BL than those in TBL and TL. However, total GPP and ET estimates are smaller in BL than in TBL and TL according to Figures 2-2, 2-3 and 2-4, indicating that the underestimation of BL are mainly attributed to the underestimation of GPP and ET from light-unsaturated leaves. Light-unsaturated leaves includes shaded leaves and sunlit leaves with low solar irradiance. Because high LAI often indicates high percentage for shaded leaves, the underestimation of fluxes estimates in BL increases with LAI.
2.4.4 Difference between TBL and TL

*TBL* and *TL* both apply the two-leaf radiation regime, so the differences in their ET and GPP estimates are not induced by the simulation of radiation. Though *TBL* tends to underestimate total ET and overestimate total GPP relative to *TL*, we find that the sunlit and shaded parts of the canopies are affected differently using *TBL* (Figure 2-6).

![Figure 2-6](image)

**Figure 2-6.** The difference between the ET estimates (a) and GPP estimates (b) from *TBL* and *TL* for sunlit leaves and shaded leaves. Negative values mean *TBL* underestimates fluxes relatively to *TL*; positive values mean *TBL* overestimates fluxes relatively to *TL*.

Figure 2-6 demonstrates that for sunlit leaves, ET estimated by *TBL* and *TL* are similar to each other, while sunlit GPP is underestimated by *TBL* relative to estimates from *TL*. For shaded leaves, *TBL* underestimates ET at all sites with the ET underestimation increases with LAI. *TBL* overestimates shaded GPP at 5 sites and underestimates at 4 sites, and the difference between the GPP estimates from *TBL* and *TL* display significant correlation with LAI. The difference between the estimated GPP and ET for shaded leaves is more pronounced than that for sunlit leaves.

In order to identify the reasons for the different estimates between *TBL* and *TL*, the simulation of ET is expressed in the form of diffusion equations:
\[ ET_{TBL,j} = G_s,j \times (e_s-e_a) \]  
(2-10)

\[ ET_{TL,j} = ET_{leaf,j} \times LAI_j = g_s,j \times (e_s-e_a) \times LAI_j \]  
(2-11)

and for GPP simulation these equations are:

\[ GPP_{TBL,j} = G_s,j \times (C_a-C_i) \]  
(2-12)

\[ GPP_{TL,j} = GPP_{leaf,j} \times LAI_j = g_s,j \times (C_a-C_i) \times LAI_j \]  
(2-13)

where \( j \) refers to sunlit or shaded leaves, \( e_a \) is the atmospheric water vapour pressure, \( e_s \) is the saturated water pressure in plant cells, \( C_a \) is the atmospheric CO\textsubscript{2} concentration and \( C_i \) is the intercellular CO\textsubscript{2} concentration.

According to the equation (10) and (11), the difference between the ET estimates by \( TBL \) and \( TL \) is driven by the difference between \( g_c \) and the value of \( g_s \times LAI \). Figure 7 compares \( G_s \) values of \( TBL \) for sunlit and shaded leaves with the corresponding \( g_s \times LAI \) values of \( TL \) for all sites.
Figure 2-7. Comparison between the average daytime $G_s$ obtained from $TBL$ and the $g_s \times LAI$ obtained from $TL$ for sunlit and shaded leaves.

Figure 2-7 shows that $G_s$ simulated by $TBL$ is smaller than $g_s \times LAI$ obtained from $TL$. Shaded leaves generally show larger gaps between $TBL$ $G_s$ and $TL$ $g_s \times LAI$ than sunlit leaves. Assuming that leaves work in parallel to transport water and carbon molecules, the difference between the $G_s$ from $TBL$ and the $g_s \times LAI$ from $TL$ suggests a potential caveat in the process of calculating $G_s$ in $TBL$. In addition, because the carbon and water fluxes are coupled via $G_s$, the underestimation of $G_s$ in $TBL$ is expected to lead to an underestimation of GPP and ET relative to $TL$. 

64
However, Figure 2-6 has shown that TBL only underestimates ET for shaded leaves and GPP for sunlit leaves, while shaded GPP are sometimes even overestimated by TBL. This conflict indicates that an uncertainty in the calculation that there is another factor that drives the difference between TBL and TL for GPP estimation. Based on equation (12) and (13), we expect the $C_i$ values estimated by TL and TBL are different (Figure 2-8).

![Figure 2-8. Comparison between the daytime average $C_i$ obtained from TBL and TL for sunlit and shaded leaves.](image)

Figure 2-8 shows that $C_i$ simulated by TBL is smaller than that of TL for both sunlit and shaded leaves. The smaller $C_i$ in TBL leads to a greater gradient to drive the CO$_2$ to
diffuse from the atmosphere to the inside of leaves, and consequently compensates for the underestimation of \( G_s \) in \( TBL \) for GPP estimations. In addition, the underestimation of \( C_i \) by \( TBL \) is usually stronger at sites with large LAI values (e.g. CaCa3, CaTp3 and CaTp4), and thus, this compensation effect at these sites is strong enough to even incur overestimations of GPP by \( TBL \) (Figure 2-4).

2.5 Discussion

For the first time, our results demonstrate the differences between \( TBL \) and \( TL \) in estimating biosphere-atmosphere carbon and water exchanges. Section 2.4.4 has found the underestimation of \( G_s \) and \( C_i \) in \( TBL \) are responsible for the difference between \( TBL \) and \( TL \). The structure of \( TL \) and \( TBL \) is briefly demonstrated in Figure 2-9 to explore the driver for the underestimation of \( G_s \) and \( C_i \) for \( TBL \).

**Figure 2-9.** A schematic description of the difference between \( TL \) and \( TBL \) models. The non-linear processes in models determines that the product of the \( g_s \) from \( TL \) and LAI does not equal to \( G_s \) from \( TBL \). This diagram applies to either sunlit or shaded leaf groups.

Process-based TBMs usually consider various linear and non-linear biochemical and biophysical processes in simulating GPP and ET (Figure 2-9). In \( TL \), all these processes are performed at the leaf-level, for sunlit and shaded leaves, then the estimated fluxes of leaves are upscaled to the canopy-scale by multiplying by LAI. In \( TBL \), all these processes are simulated at the canopy-scale through upscaling the key biochemical and
biophysical parameters from leaf to canopy. If all the processes in TBM were linear, then \( TL \) would be equivalent to \( TBL \). However, due to Jenson’s inequality, \( G \), cannot be expressed as a linear function of \( g_s \) and LAI.

\( C_i \) is dynamically adjusted in the processes shown in Figure 9 until the model realizes an optimal water use efficiency [WUE; Sellers 1997; Medlyn et al, 2011; Wang et al., 2017]. With an artificially upscaled \( V_{\text{max}25} \) for a big leaf, the big leaf is apparently more capable of assimilating \( CO_2 \) compared to the leaves in reality in the same environment, and thus driving \( C_i \) to be lower in the big leaf. With the change of photosynthesis by using the big leaf, the WUE is expected to be adjusted accordingly to obtain an optimal WUE for the big leaf. This WUE is different from the WUE acquired directly at leaf level simulations.

![Graph showing WUE for sunlit and shaded leaves estimated by TBL and TL.](image)

**Figure 2-10.** WUE for sunlit and shaded leaves estimated by \( TBL \) and \( TL \).

Figure 2-10 shows that the WUE acquired from \( TBL \) is similar to that from TL for sunlit leaves, while for shaded leaves the WUE acquired from \( TBL \) is larger than its counterpart in \( TL \). Based on the separation of sunlit and shaded LAI described by equation (3) and (4), we know that the sunlit LAI could not be larger than 2 and the remaining LAI are shaded leaves. Therefore, the upscaling of \( V_{\text{max}25} \) for shaded leaves is stronger than the upscaling of \( V_{\text{max}25} \) for sunlit leaves when \( TBL \) is used. The WUE estimated by \( TBL \) also consistently increases with LAI, which red flags the analysis of WUE trend in the context of climate change using \( TBL \).
At last, considering that most biochemical and biophysical processes in TBMs are developed based on leaf-level measurements, we suggest that it is conceptually correct to apply TL to TBMs which uses $g_s$ in the mathematical formulations of these processes and avoids the uncertainties propagated from the derivation of $G$, and other canopy-level photosynthetic parameters as intermediate variables. Though the results from TL models may not be superior to the results from TBL due a range of reasons such as observational uncertainty of inputs, uncertainty of flux measurements and the uncertainty of leaf-level parameters, the difference between the estimates from TBL and TL are still worth noting since estimates from TBL show systematically increasing bias with LAI. With a given set of input parameters, the systematic differences between TL and TBL models at all test sites suggest that more attention should be given to model structure in addition to improving model parameters. In fact, using a model with correct structure and processes should be a prerequisite to tuning model parameters in the quest to understand the complex processes governing the carbon and water fluxes of terrestrial ecosystems.

2.6 Conclusion

The “big-leaf” concept is widely used to describe the bulk control of plant canopies on transporting water and carbon molecules. It is characterized by the use of canopy stomatal conductance ($G_s$) in the Penman-Monteith equation in a two-layers (i.e. vegetation and soil) frame. In order to consider the physiological principle of carbon-water coupling, some state-of-the-art TBMs expand the “big-leaf” concept by upscaling leaf-level photosynthetic parameters to their canopy-level counterparts, and directly using leaf-level biochemical models at the canopy scale. $G_s$ is then calculated in BL for ET simulation. However, BL has been reported to incur some biases in GPP estimation, and TBL has been developed to address the problem [e.g., De Pury and Farquhar, 1997]. Meanwhile, less attention has been paid to the uncertainties underlying the artificial upscaling process for $G_s$ and other biochemical parameters. In this study, we aim to promote the use of TL in TBMs built on the carbon-water coupling principle and to avoid the use of $G_s$ in the Penman-Monteith equation for ET and GPP modelling. The
performance of BL, TBL and TL in estimating ET and GPP is evaluated with flux measurements from nine eddy-covariance towers. Our conclusions are as follows:

(1) BL underestimates ET and GPP across all sites because Beer’s law fails to correctly describe the instantaneous radiation distribution in canopies. Increasing LAI leads to growing underestimations of ET and GPP in BL, mainly due to the increasing underestimation of fluxes from shaded leaves. GPP and ET estimates from BL are 25% and 16% lower than the estimates from TBL, respectively.

(2) TBL estimates ET 7% lower relative to TL across nine sites, mainly due to the underestimation of ET from shaded leaves. TBL and TL produce very similar total GPP values when LAI is low. The partition of sunlit and shaded component of GPP is different to the partition in TL. The difference between the fluxes estimated by TBL and TL is due to the underestimation of $G_s$ and $C_i$ in TBL, and the difference in ET and GPP between TL and TBL aggravates with increasing LAI.

(3) The inadequate performance of BL and TBL invalidates the “big-leaf” assumption underlying the PM equation in carbon-water coupled TBMs. The non-linear biophysical and biochemical processes make BL and TBL suffer from Jenson’s inequality, which drives the artificial upscaling of leaf-level parameters to canopy-scale parameters and leads to incorrect simulations of $G_s$ and $C_i$. TL is physically correct since it couples the water flow with the carbon flow at the leaf level by directly using the leaf stomatal conductance derived from leaf biochemical models in the PM equation. We therefore suggest TL as the optimal upscaling scheme in carbon-water coupled TBMs.

2.7 Appendix

2.7.1 Nitrogen-weighted $V_{\text{max}}^{25}$ and $J_{\text{max}}^{25}$ for sunlit and shaded leaves

Chen et al. (2012) combined the “two-leaf” separation scheme and a nitrogen gradient to derive the $V_{\text{max}}^{25}$ and $J_{\text{max}}^{25}$ values for the sunlit and shaded leaves separately. Leaf nitrogen
content per leaf area $N(L)$ generally decreases exponentially from the top to the bottom in a canopy (Equation A1):

$$N(L) = N_0 e^{-k_n L} \quad (2-A1)$$

where the extinction coefficient $k_n = 0.3$ used in BEPS is adopted from De Pury and Farquhar (1997), $N_0$ is the nitrogen content at top of the canopy, and $L$ is the canopy depth described in total LAI. On the other hand, the leaf maximum carboxylation rate at 25 °C ($V_{max}^{25}$) is proportional to the leaf nitrogen content therefore it can be expressed as:

$$V_{max}^{25}(L) = V_{max \_0}^{25} \chi_n N(L) \quad (2-A2)$$

where $V_{max \_0}^{25}$ is the $V_{max}^{25}$ of the leaves at the top of the canopy, $\chi_n$ quantifies the relative change of $V_{max}^{25}$ to the leaf nitrogen content in the canopy. $\chi_n$ has units of m²/g while $N(L)$ has units of g/m². The value of $\chi_n$, the mean value of $N$ and its standard deviation, and the standard deviation of $V_{max}^{25}$ are provided according to the plant functional types [Chen et al., 2012]. $N_0$ is taken as the mean N value plus one standard deviation, $V_{max \_0}^{25}$ is taken as the input $V_{max}^{25}$ value plus one standard deviation.

The fraction of the sunlit and shaded leaves in the canopy change with the canopy depth:

$$f_{sun}(L) = \Omega e^{-kL} \quad (2-A3)$$

$$f_{sh}(L) = 1 - \Omega e^{-kL} \quad (2-A4)$$

where $k = G(\theta)\Omega / \cos \theta$. $G(\theta)$ is the projection coefficient of the canopy and it is 0.5 assuming a spherical leaf angle distribution. $\Omega$ is the clumping index, and $\theta$ is the solar zenith angle. We assume the $V_{max}^{25}$ of a representative sunlit or shaded leaf is equal to the mean $V_{max}^{25}$ value of the sunlit or shaded leaves’ group. Therefore, the $V_{max}^{25}$ of a representative sunlit or shaded leaf is obtained by the following integrations:

$$V_{max \_sunlit}^{25} = \frac{\int_0^L V_{max \_0}^{25} \chi_n N(L) f_{sun}(L) dL}{\int_0^L f_{sun}(L) dL} = V_{max \_0}^{25} \chi_n N_0 \frac{\int_0^L e^{-k_n L} \Omega e^{-kL} dL}{\int_0^L \Omega e^{-kL} dL} =$$

$$V_{max \_0}^{25} \chi_n N_0 \frac{k[1-e^{-(k_n+k)L}]}{(k_n+k)(1-e^{-kL})} \quad (2-A5)$$
After the $V_{\text{max \_shaded}}^{25}$ values of the representative sunlit and shaded leaves are obtained, the maximum electron transport rate at 25 °C ($J_{\text{max \_25}}$) is obtained using the following equations [Medlyn et al., 1999]

\[
J_{\text{max \_sunlit}}^{25} = 2.39V_{\text{max \_sunlit}}^{25} - 14.2 \quad (2-A7)
\]

\[
J_{\text{max \_shaded}}^{25} = 2.39V_{\text{max \_shaded}}^{25} - 14.2 \quad (2-A8)
\]

### 2.7.2 Leaf energy budget

In the absence of rainfall and snow coverage over leaves, the leaf energy budget is composed of the net radiation on leaf ($R_n$), the sensible heat ($Q$) and the latent heat ($LE$) from the leaf in every hourly step, during which period the heat storage of leaf is negligible.

\[
R_n = H + LE \quad (2-B1)
\]

#### 2.7.2.1 Net radiation on a leaf

In BEPS the whole canopy was divided into four groups of leaves based on the location and radiation features of the leaves, namely sunlit leaves in the overstorey, shaded leaves in the overstorey, sunlit leaves in the understorey and shaded leaves in the understorey [Chen et al., 1999; Liu et al., 2003]. The leaves in each group have identical features so BEPS could use one leaf to represent a group. Net radiation on a leaf comprises three sources:

\[
R_{n,i} = R_{\text{dir},i} + R_{\text{dif},i} + R_{l,i} \quad (2-B2)
\]
where \( R_n \) is the total net radiation on a given leaf, \( R_{dir}, R_{dif} \) and \( R_i \) refers to the net direct incoming solar radiation, net diffuse solar radiation and net longwave radiation on the leaf. \( i \) refers to one of the four types of leaves. For a shaded leaf, \( R_{dir} = 0 \).

In order to differentiate the incoming solar radiation into a direct and diffuse part, a semi-empirical equation is applied:

\[
\frac{S_{dif}}{S_g} = \begin{cases} 
0.943 + 0.734r - 4.9r^2 + 1.796r^3 + 2.058r^4 & r < 0.8 \\
0.13 & r \geq 0.8
\end{cases} \tag{2-B3}
\]

\[
S_{dir} = S_g - S_{dif} \tag{2-B4}
\]

where \( S_g, S_{dir} \) and \( S_{dif} \) are incident solar irradiance, incoming direct solar radiation and diffuse solar radiation, respectively. \( r \) is a parameter used to quantify the cloudiness of the sky,

\[
r = \frac{S_g}{S_0 \cos \theta} \tag{2-B5}
\]

where \( S_0 \) is the solar constant set as 1362 W/m\(^2\), \( \theta \) is the solar zenith angle.

The net direct solar radiation on the sunlit representative leaf in the overstorey or understorey of the canopy is:

\[
R_{dir \_sunlit} = R_{dir \_u \_sunlit} = (1 - \alpha_L)S_{dir \cos \alpha / \cos \theta} \tag{2-B6}
\]

where \( \alpha_L \) is the albedo of the leaves. But in BEPS, \( \alpha_L \) is different for the overstorey and the understorey because snow coverage varies with canopy depth. \( \alpha \) is the mean leaf-sun angle which is fixed at 60° when the canopy has a spherical leaf distribution.

On the other hand, the net diffuse solar radiation on the four groups of the leaves are approximated respectively as:

\[
R_{dif \_sunlit} = R_{dif \_o \_sunlit} = (1 - \alpha_L)(S_{dif \left[ 1 - e^{-0.5kLAI_o / \cos \theta} \right]} / LAI_o + C_o) \tag{2-B7}
\]
\[ R_{dif \_u \_sunlit} = R_{dif \_u \_shaded} = (1 - \alpha_L) \left( S_{dif} e^{-0.5 \Delta L A I_u / \cos \tilde{\theta}_o} \left[ 1 - e^{-0.5 \Delta L A I_u / \cos \tilde{\theta}_u} \right] / \Delta L A I_u + C_u \right) \]  

(2-B8)

where \( \Delta L A I_o \) and \( \Delta L A I_u \) are the LAI value of the overstorey and the understorey, \( C_o \) and \( C_u \) are used to quantify the multiple scattering of the direct solar radiation from the leaf (Chen et al., 1999))

\[ C_o = 0.07 \Omega S_{dir}(1.1 - 0.1 \Delta L A I) e^{-\cos \theta} \]  

(2-B9)

\[ C_u = 0.07 \Omega S_{dir} e^{-0.5 \Delta L A I_o / \cos \theta}(1.1 - 0.1 \Delta L A I_u) e^{-\cos \theta} \]  

(2-B10)

\( \tilde{\theta}_o \) and \( \tilde{\theta}_u \) are the representative zenith angles for diffuse radiation transmission of the overstorey and understory leaves and slightly dependent on the corresponding LAI [Liu et al., 2003]:

\[ \cos \tilde{\theta} = 0.537 + 0.025 \Delta L A I \]  

(2-B11)

The net longwave radiation on these leaves is calculated as:

\[ R_{L \_o \_sunlit} = R_{L \_o \_shaded} = \frac{1}{\Delta L A I_o} \left\{ \left[ \varepsilon_o \left[ \varepsilon_a \sigma T_a^4 + \varepsilon_u \sigma T_u^4 \left( 1 - e^{-0.5 \Delta L A I_u \Omega / \cos \tilde{\theta}_u} \right) \right] + \varepsilon_g \sigma T_g^4 \right] e^{-0.5 \Delta L A I_o \Omega / \cos \tilde{\theta}_o} \right\} \]  

(2-B12)

\[ R_{L \_u \_sunlit} = R_{L \_u \_shaded} = \frac{1}{\Delta L A I_u} \left\{ \left[ \varepsilon_u \left[ \varepsilon_a \sigma T_a^4 e^{-0.5 \Delta L A I_o \Omega / \cos \tilde{\theta}_o} + \varepsilon_o \sigma T_o^4 \right] \right] e^{-0.5 \Delta L A I_u \Omega / \cos \tilde{\theta}_u} \right\} \]  

(2-B13)

where \( \sigma \) is the Stephen-Boltzmann constant equals to \( 5.67 \times 10^{-8} \text{W} \text{m}^{-2} \text{K}^{-4} \). \( \varepsilon_a, \varepsilon_o, \varepsilon_u \) and \( \varepsilon_g \) are the emissivity of the atmosphere, overstorey, understory and ground surface, respectively. \( \varepsilon_o, \varepsilon_u \) and \( \varepsilon_g \) are prescribed as 0.98, 0.98 and 0.95, respectively, according
to [Chen and Zhang, 1989; Chen et al., 1989], and \( e_a \) is computed as 
\[
\epsilon_a = 1.24 \left( \frac{e_a}{T_a} \right)^{1/7} \]
[Brutsaert, 1982], where \( e_a \) and \( T_a \) are water vapor pressure in mbar and temperature of the atmosphere in K. \( T_o, T_u \) and \( T_g \) are the temperatures of the overstorey, the understorey and ground, respectively, in K, and \( T_o \) and \( T_u \) are calculated as the weighed average temperature of sunlit leaves and shaded leaves in overstorey and understorey, respectively.

### 2.7.2.2 Sensible heat from a leaf

The sensible heat is calculated for overstorey sunlit leaves, overstorey shaded leaves, understorey sunlit leaves and understorey shaded leaves, respectively.

\[
Q_i = (T_i - T_a) \rho c_p g_{Hi} \tag{2-B14}
\]

where \( i \) refers to the type of the leaf, \( \rho \) is the density of air, \( c_p \) is the specific heat of air, \( g_H \) is total conductance of heat from the leaf surface to the atmosphere, which equals to the reciprocal of the leaf boundary layer resistance and aerodynamic resistance in tandem.

### 2.7.2.3 Latent heat from a leaf

Latent heat is calculated using the Penman-Monteith equation (Equation 5), which was simplified into a linear function of leaf temperature [Campbell and Norman, 2012] in BEPS:

\[
LE_i = \left( VPD + \Delta(T_i - T_a) \right) \rho c_p g_{W,i}/\gamma \tag{2-B15}
\]

where \( i, \rho \) and \( c_p \) have the same meaning as above, \( VPD \) is the vapor pressure deficit of the ambient air, \( \gamma \) is the psychrometric constant, \( \Delta \) is the slope of the saturation vapor pressure curve at air temperature, \( g_w \) is total conductance of water vapor from leaf interior to the atmosphere, which equals to the reciprocal of the tandem of the leaf boundary layer resistance, aerodynamic resistance and leaf stomatal resistance (1/\( g_s \)). \( g_s \) is obtained from the carbon assimilation module using Farquhar’s model and the Ball-Woodrow-Berry equation.
Ultimately, the three components of leaf energy budget are expressed as a function of leaf temperature. We reiterate the processes above until the leaf temperature converge to realize the leaf energy balance.

### 2.7.3 Quantification of the soil water stress factor

To account for the effect of the soil water deficit on stomatal conductance, a soil water stress factor ($f_w$) based on the ratio of the measured available water in the soil to the maximum plant available water [Wigmosta et al., 1994; Wang and Leuning, 1998; Chen et al., 2005] was calculated as follows:

$$ f_w = \begin{cases} 0 & \theta_{sw}(z) - \theta_{wp} < 0 \\ \frac{\theta_{wp} - \theta_{sw}(z) - \theta_{fc}}{\theta_{wp} - \theta_{sw}(z)} & 0 \leq \theta_{sw}(z) \leq \theta_{fc} \\ 1 & \theta_{sw}(z) > \theta_{fc} \end{cases} \quad (2-C1) $$

where $\theta_{sw}(z)$ is the soil water content of layer $z$, $z$ often refers to the top 30 cm based on the availability of the soil water measurements. $\theta_{wp}$ and $\theta_{fc}$ are the wilting point and the field capacity, respectively, (m$^3$/m$^3$) of the soil layer. $\theta_{wp}$ and $\theta_{fc}$ are derived from the soil texture information provided by Fluxnet (http://fluxnet.ornl.gov/), the patterns of multi-year soil moisture measurements and the algorithm developed by Saxton and Rawls (2006).

### 2.7.4 Parameterization for TBL

According to literature [Wang and Leuning, 1998; Dai et al., 2004; Ryu et al., 2011], TBL will upscale the leaf-level $V_{max}^{25}$ to its canopy counterpart first, then it will calculate $A_c$ and $g_c$ directly without the derivation of the parameter $A$ and $g_s$. In this case,

$$ V_{max, canopy}^{25} = \int_{0}^{L} V_{max, 0}^{25} \chi N(L) dL \quad (2-D1) $$

$$ V_{max, sunlit, canopy}^{25} = \int_{0}^{L} V_{max, 0}^{25} \chi N(L) f_{sun}(L) dL \quad (2-D2) $$
\[ V_{\text{max\_shaded\_canopy}}^{25} = V_{\text{max\_canopy}}^{25} - V_{\text{max\_sunlit\_canopy}}^{25} = \int_0^L V_{\text{max\_0}}^{25} \chi_n N(L) f_{\text{sh\_}}(L) dL \]

(2-D3)

where \( V_{\text{max\_canopy}}^{25} \), \( V_{\text{max\_sunlit\_canopy}}^{25} \) and \( V_{\text{max\_shaded\_canopy}}^{25} \) are the canopy-level \( V_{\text{max}}^{25} \) for the whole canopy, sunlit canopy and shaded canopy, respectively. \( N(L) \) is the nitrogen gradient in canopy, and \( f_{\text{sun\_}}(L) \) and \( f_{\text{sh\_}}(L) \) are the fraction of sunlit and shaded leaves in the canopy that change with the canopy depth (Appendix 2.7.1).

Through using the canopy-scale \( V_{\text{max}}^{25} \) in Farquhar’s biochemical model and the Ball-Woodrow-Berry stomatal conductance model, we obtain the \( g_c \) and \( A_c \) for the sunlit canopy and shaded canopy, respectively.
2.7.5 Correlations between the simulated hourly ET (GPP) and measured hourly ET (GPP) under all schemes.

Table 2-A1 shows that on average, simulations using BL, TBL and TL explains about 67%, 70% and 71% of the variance in the ET measurements, respectively. Linear correlations between the simulations and the measurements indicate that TL performs best in capturing the temporal patterns of ET with a regression slope of 0.91, while BL and TBL underestimate ET with slopes of 0.72 and 0.83, respectively. Average RMSEs between simulated and measured ET are 0.055, 0.055 and 0.051 mm/h using BL, TBL and TL, respectively.

In the linear regressions between simulated GPP and measured GPP, the mean $r^2$ values are 0.69, 0.81 and 0.82, and the mean slopes are 0.66, 0.95 and 0.92 for BL, TBL, and TL, respectively. Moreover, the mean RMSEs are 0.135, 0.112 and 0.107 g/m²/h for BL, TBL, and TL, respectively. TBL and TL simulate GPP with similar accuracies, while BL significantly underestimates GPP. The variations of these statistics across the sites are smaller for TL or TBL than for BL, suggesting that TL or TBL is more suitable for large-scale applications.

**Table 2-A1. Correlations between the simulated hourly ET (GPP) and measured hourly ET (GPP) under all schemes**

<table>
<thead>
<tr>
<th>Site ID</th>
<th>year</th>
<th>Upscaling schemes</th>
<th>ET $r^2$</th>
<th>Slope</th>
<th>Intercept (mm/h)</th>
<th>RMSE (mm/h)</th>
<th>GPP $r^2$</th>
<th>Slope</th>
<th>Intercept (g/m²/h)</th>
<th>RMSE (g/m²/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaCa2</td>
<td>2002-2010</td>
<td>BL</td>
<td>0.71</td>
<td>0.76</td>
<td>0.006</td>
<td>0.04</td>
<td>0.72</td>
<td>0.015</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBL</td>
<td>0.73</td>
<td>0.89</td>
<td>0.003</td>
<td>0.04</td>
<td>0.84</td>
<td>0.009</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TL</td>
<td>0.74</td>
<td>0.92</td>
<td>0.004</td>
<td>0.04</td>
<td>0.85</td>
<td>0.015</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>CaCa3</td>
<td>2002-2009</td>
<td>BL</td>
<td>0.69</td>
<td>0.71</td>
<td>0.007</td>
<td>0.05</td>
<td>0.72</td>
<td>0.018</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBL</td>
<td>0.72</td>
<td>0.84</td>
<td>0.004</td>
<td>0.05</td>
<td>0.84</td>
<td>0.013</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TL</td>
<td>0.76</td>
<td>0.98</td>
<td>0.005</td>
<td>0.04</td>
<td>0.84</td>
<td>0.016</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>CaCbo</td>
<td>2008-2013</td>
<td>BL</td>
<td>0.55</td>
<td>0.52</td>
<td>0.025</td>
<td>0.09</td>
<td>0.65</td>
<td>0.047</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBL</td>
<td>0.65</td>
<td>0.62</td>
<td>0.021</td>
<td>0.08</td>
<td>0.77</td>
<td>0.052</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TL</td>
<td>0.65</td>
<td>0.71</td>
<td>0.023</td>
<td>0.08</td>
<td>0.78</td>
<td>0.054</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>CaGro</td>
<td>2005-2011</td>
<td>BL</td>
<td>0.69</td>
<td>0.76</td>
<td>0.016</td>
<td>0.05</td>
<td>0.77</td>
<td>0.011</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBL</td>
<td>0.72</td>
<td>0.89</td>
<td>0.011</td>
<td>0.06</td>
<td>0.82</td>
<td>0.007</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TL</td>
<td>0.73</td>
<td>0.94</td>
<td>0.011</td>
<td>0.05</td>
<td>0.84</td>
<td>0.009</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>CaOas</td>
<td>2002-2010</td>
<td>BL</td>
<td>0.73</td>
<td>0.80</td>
<td>0.013</td>
<td>0.05</td>
<td>0.82</td>
<td>0.017</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBL</td>
<td>0.79</td>
<td>0.89</td>
<td>0.010</td>
<td>0.05</td>
<td>0.90</td>
<td>0.006</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TL</td>
<td>0.80</td>
<td>0.91</td>
<td>0.009</td>
<td>0.05</td>
<td>0.90</td>
<td>0.011</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL</td>
<td>TBL</td>
<td>TL</td>
<td>BL</td>
<td>TBL</td>
<td>TL</td>
<td>BL</td>
<td>TBL</td>
<td>TL</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>CaObs</td>
<td>2002-2009</td>
<td>0.62</td>
<td>0.89</td>
<td>0.015</td>
<td>0.06</td>
<td>0.61</td>
<td>0.68</td>
<td>0.006</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.62</td>
<td>1.03</td>
<td>0.012</td>
<td>0.06</td>
<td>0.67</td>
<td>0.99</td>
<td>-0.001</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.65</td>
<td>1.08</td>
<td>0.012</td>
<td>0.06</td>
<td>0.69</td>
<td>0.96</td>
<td>0.005</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>CaOjp</td>
<td>2005-2010</td>
<td>0.57</td>
<td>0.67</td>
<td>0.020</td>
<td>0.04</td>
<td>0.69</td>
<td>0.63</td>
<td>0.011</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.57</td>
<td>0.74</td>
<td>0.019</td>
<td>0.05</td>
<td>0.67</td>
<td>0.82</td>
<td>0.008</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.57</td>
<td>0.76</td>
<td>0.020</td>
<td>0.05</td>
<td>0.72</td>
<td>0.85</td>
<td>0.013</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>CaTp3</td>
<td>2009-2013</td>
<td>0.74</td>
<td>0.65</td>
<td>0.016</td>
<td>0.05</td>
<td>0.82</td>
<td>0.54</td>
<td>0.034</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.77</td>
<td>0.78</td>
<td>0.011</td>
<td>0.04</td>
<td>0.88</td>
<td>0.87</td>
<td>0.046</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.79</td>
<td>0.90</td>
<td>0.012</td>
<td>0.04</td>
<td>0.88</td>
<td>0.80</td>
<td>0.045</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>CaTp4</td>
<td>2008-2013</td>
<td>0.71</td>
<td>0.68</td>
<td>0.010</td>
<td>0.06</td>
<td>0.82</td>
<td>0.60</td>
<td>0.019</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.73</td>
<td>0.75</td>
<td>0.007</td>
<td>0.06</td>
<td>0.89</td>
<td>0.98</td>
<td>0.027</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.74</td>
<td>0.95</td>
<td>0.006</td>
<td>0.06</td>
<td>0.89</td>
<td>0.87</td>
<td>0.025</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>
2.8 References


Chen, J. M., B. J. Yang, and R. H. Zhang (1989), Soil thermal emissivity as affected by its water content and surface treatment, *Soil science (USA).*


Cramer, W. et al. (2001), Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models,


Ryu, Y. et al. (2011), Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales, *Global Biogeochemical Cycles*, 25(4), n/a-n/a, doi:10.1029/2011GB004053.


INTRODUCTION.


Chapter 3

Incorporating leaf chlorophyll content into a two-leaf terrestrial biosphere model for estimating carbon and water fluxes at a forest site

Preamble: this chapter is accepted for publication in Agriculture and Forestry Meteorology. The authors of this paper are: X. Luo, H. Croft, J. M. Chen, P. Bartlett, R. Staebler and N. Froelich. XL performed the data analysis and wrote the paper, HC and XL collected the field data, XL, HC and JC collectively developed the idea. PB provided equipment to support the field work, RS and NF processed and provided the flux data and ancillary data. All authors contributed to the writing.

3.1 Abstract

Chlorophyll is the main light-harvesting pigment in leaves, facilitating photosynthesis and indicating the supply of nitrogen for photosynthetic enzymes. In this study, we explore the feasibility of integrating leaf chlorophyll content (Chl_{leaf}) into a Terrestrial Biosphere Model (TBM), as a proxy for the leaf maximum carboxylation rate at 25°C ($V_{max}^{25}$), for the purpose of improving carbon and water flux estimation. Measurements of Chl_{leaf} and $V_{max}^{25}$ were made in a deciduous forest stand at the Borden Forest Research Station in southern Ontario, Canada, where carbon and water fluxes were measured by the eddy covariance method. The use of Chl_{leaf}-based $V_{max}^{25}$ in the TBM significantly reduces the bias of estimated gross primary productivity (GPP) and evapotranspiration (ET) and improves the temporal correlations between the simulated and the measured fluxes, relative to the commonly employed cases of using specified constant $V_{max}^{25}$, leaf area index (LAI)-based $V_{max}^{25}$ or specific leaf area (SLA)-based $V_{max}^{25}$. The biggest improvements are found in spring and fall, with the mean absolute errors (MAEs) between modelled and measured GPP reduced from between 2.2-3.2 to 1.8 g C m$^{-2}$ d$^{-1}$ in spring and from between 2.1-2.8 to 1.8 g C m$^{-2}$ d$^{-1}$ in fall. The MAEs in ET estimates are reduced from 0.7-0.8 mm d$^{-1}$ to
0.6 mm d⁻¹ in spring, but no significant improvement is noted in autumn. A two-leaf upscaling scheme is used to account for the uneven distribution of incoming solar radiation inside canopies and the associated physiological differences between leaves. We find that modelled $V_{\text{max}}^{25}$ in sunlit leaves is 34% larger than in the shaded leaves of the same Chl$_{\text{leaf}}$, which echoes previous physiological studies on light acclimation of plants. This study represents the first case of the incorporation of chlorophyll as a proxy for $V_{\text{max}}^{25}$ in a two-leaf TBM at a forest stand and demonstrates the efficacy of using chlorophyll to constrain $V_{\text{max}}^{25}$ and reduce the uncertainties in GPP and ET simulations.

3.2. Introduction

Predicting gross primary productivity (GPP) and evapotranspiration (ET) has posed a challenge for accurately quantifying the global terrestrial carbon, water and energy budgets in the context of climate change [Jung et al., 2010; Schaefer et al., 2012]. State-of-the-art terrestrial biosphere models (TBMs) have been developed to quantify carbon and water fluxes by describing physical and physiological processes in the soil-vegetation-atmosphere continuum. Model validation and inter-model comparisons have suggested that TBMs are particularly sensitive to the parameterization of leaf photosynthetic capacity [Kattge et al., 2009]. Since leaf photosynthetic parameters can only be measured directly at the leaf level [Wilson et al., 2000; Xu and Baldocchi, 2003], using labour-intensive instruments, their limited availability in both space and time has hindered progress in improving GPP and ET estimates by TBMs. Consequently, finding a reliable and easily measurable proxy for leaf photosynthetic parameters across large spatial scales is paramount for the modelling community.

Most TBMs simulate photosynthesis using an enzyme-kinetic model developed by Farquhar [Farquhar et al., 1980]. Two key parameters are used to represent the leaf photosynthetic capacity: the maximum carboxylation capacity ($V_{\text{max}}^{25}$) and the maximum electron transport capacity ($J_{\text{max}}^{25}$) at 25°C. $V_{\text{max}}^{25}$ refers to the kinetic properties of the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) within the Calvin-Benson cycle. $J_{\text{max}}^{25}$ is related to the intrinsic properties of the thylakoid membrane, which
corresponds to the availability of the cytochrome b6f complex (Cty f) [Hikosaka and Terashima, 1995] to transport electrons in order to produce chemical energy transfers such as ATP and NADPH. ATP and NADPH are then used to drive the Calvin-Benson Cycle by reducing Rubisco into ribulose bisphosphate (RuBP). On average about two electrons are needed to reduce one unit of Rubisco, implying a constant ratio between \( J_{\text{max}}^{25} \) and \( V_{\text{max}}^{25} \). Frequently, \( V_{\text{max}}^{25} \) and \( J_{\text{max}}^{25} \) are each prescribed as a constant value for a given plant functional type (PFT) based on limited measurements [Wullschleger, 1993; Medlyn et al., 1999]. Owing to the fact that Rubisco and Cty f are both rich in nitrogen, some studies derived \( V_{\text{max}}^{25} \) and \( J_{\text{max}}^{25} \) from a global compilation of nitrogen traits in plants, assuming a linear relationship between Rubisco and leaf nitrogen content per unit area (\( N_{\text{area}} \)) [Kattge et al., 2009; Walker et al., 2014]. However, accurately representing temporal and spatial variations in both \( V_{\text{max}}^{25} \) and \( J_{\text{max}}^{25} \) is still elusive, as a robust observational methodology to quantify the dynamics of \( N_{\text{area}} \) across large spatial extents is lacking [Knyazikhin et al., 2013]. Moreover, some studies have questioned the efficacy of the nitrogen-based method because the relationship between \( N_{\text{area}} \) and Rubisco varies temporally due to the dynamic allocation of \( N_{\text{area}} \) between photosynthetic and non-photosynthetic components in leaves, particularly over a growing season [Croft et al., 2017; Kalacska et al., 2015].

Recent research has drawn attention to the potential of using leaf chlorophyll content (\( \text{Chl}_{\text{leaf}} \)) as an alternative constraint on leaf photosynthetic capacities [Croft et al., 2015a, 2017; Houborg et al., 2013; Koffi et al., 2015; Alton, 2017]. Chlorophyll molecules located in light-harvesting-complexes (LHCs) absorb photons, which are then transported to initiate specific redox reactions of chlorophyll molecules in Photosystem II to donate electrons to the electron transport chain (ETC). The concentration of \( \text{Chl}_{\text{leaf}} \) is related to the rate of photosynthesis through its determining role in the instantaneous electron transport rate (J) [Porcar-Castell et al., 2014]. As \( \text{Chl}_{\text{leaf}} \) is a nitrogen rich pigment, \( \text{Chl}_{\text{leaf}} \), Rubisco and Cty f are linearly related to each other according to their common correlations with the leaf photosynthetic nitrogen pool [Terashima and Evans, 1988; Evans, 1989b; Hikosaka and Terashima, 1996]. Recent research has demonstrated that it is possible to model \( J_{\text{max}}^{25} \) and \( V_{\text{max}}^{25} \) from \( \text{Chl}_{\text{leaf}} \) using simple linear equations [Houborg et al., 2013; Croft et al., 2017]. This potential role of \( \text{Chl}_{\text{leaf}} \) as a proxy for photosynthetic capacity has
important implications for improved modelling at regional to global scales, through the ability of estimating Chl_{leaf} from remotely sensed data, using empirical [Croft et al., 2014a; Gitelson et al., 2005; Wu et al., 2008] and physically-based models [Croft et al., 2013; Croft et al., 2015b; Houborg et al., 2015a; Zhang et al., 2007]. Satellite-derived Chl_{leaf} could provide an accurate, achievable way of quantifying J_{max}^{25} and V_{max}^{25} in a temporally and spatially explicit manner.

Based on the assumptions above, Houborg et al. [2013] proposed a framework to incorporate measured or satellite-derived Chl_{leaf} into TBMs by building universal Chl_{leaf}–V_{max}^{25} relationships for C3 and C4 plants. Their inclusion of Chl_{leaf} into a TBM helped the GPP simulations to better capture the daily and seasonal variations of observed fluxes over a corn field. Beyond that, limited work has been done to incorporate Chl_{leaf} in TBMs, except for works that have related Chl_{leaf} to vegetation productivity using empirical light use efficiency (LUE) approaches [Gitelson et al., 2006; Houborg et al., 2011; Croft et al., 2015]. As forests account for 50% of the global GPP [Pan et al., 2013] and 45% of the global ET [Oki and Kanae, 2006], there is a growing interest in exploring the potential to use Chl_{leaf}–V_{max}^{25} relationship in forests to improve the prediction of carbon and water fluxes in these areas.

The complex structures of forest canopies present greater challenges in the upscaling process than for homogenous croplands. Canopy architecture leads to an uneven distribution of solar irradiance from the sunlit top of trees to the shaded interior of a canopy. Values of N_{area} are larger in the sunlit part of the canopy than in shaded leaves, with N_{area} declining from the top to the bottom of a canopy, mimicking the gradient of long-term radiation distribution [Field, 1983; Hirose and Werger, 1987; Niinemets, 1997; Warren and Adams, 2001]. Along with the changes in N_{area}, the abundance of the nitrogen-rich photosynthetic components -- Rubisco and Cyt f -- is reported to be highly correlated with N_{area} and varies from the top to the bottom of the canopy in proportion to N_{area} [Sage et al., 1987; Evans, 1989b]. However, Chl_{leaf} is reported to be largely unrelated to the N_{area} gradient and is almost uniform for leaves under various light environments [Anderson et al., 1988; Terashima and Evans, 1988; Iio et al., 2005; Walters, 2005; Lichtenthaler et al., 2007; Lambers et al., 2008], although some studies show a slight chlorophyll gradient
These trends lead to a substantially higher $\text{Chl}_{\text{leaf}} : \text{N}_{\text{area}}$ ratio in shaded conditions than in full sunlight, which imposes variations on the $\text{Chl}_{\text{leaf}} - V_{\text{max}}^{25}$ relationships inside a canopy. This phenomenon has been attributed to an optimal nitrogen allocation in plants, to maximise the overall productivity of the canopy [Kull, 2002]. Accordingly, for shaded leaves, relatively more nitrogen will be allocated to chlorophyll in order to harvest photons, while for sunlit leaves, relatively more nitrogen will be invested in Rubisco and Cty f to accelerate dark reactions [Evans, 1989a; Hikosaka and Terashima, 1996; Hikosaka, 2014].

In this study, a TBM incorporating a two-leaf upscaling scheme is used to estimate GPP and ET within a deciduous forest, across two growing seasons. The two-leaf scheme separates a canopy into groups -- sunlit leaves and shaded leaves -- based on the first-order features of instantaneous solar irradiance on leaves [Sinclair et al., 1976; Norman, 1982; Chen et al., 1999]. It is a robust description of the complex leaf light environment in canopies and has been proven to be more capable of simulating GPP and ET variations than the commonly used big-leaf scheme [Chen et al., 1999; De Pury and Farquhar, 1997; Luo et al., in review]. The two-leaf scheme delivers an available and potential tool to account for the variations in the $\text{Chl}_{\text{leaf}} - V_{\text{max}}^{25}$ relationship in a complex forest canopy, as well as for other physiological traits affected by the light environment. Using intensive leaf-level and canopy-level measurements at a temperate broadleaf forest, the specific objectives of this study are to:

1) investigate the improvements in GPP and ET simulations from a two-leaf TBM that uses $\text{Chl}_{\text{leaf}}$ to constrain $V_{\text{max}}^{25}$;

2) evaluate the differences in the $\text{Chl}_{\text{leaf}} - V_{\text{max}}^{25}$ relationship, and other physiological traits, between sunlit and shaded leaves.

### 3.3 Materials and methods

#### 3.3.1 Site profile
The Borden Forest Research Station is a mixed temperate forest site located in southern Ontario, Canada (44°19’N, 79°56’W) [Froelich et al., 2015]. It lies within an ecotone that extends across eastern North America between 44 and 47°N and contains both southern temperate species and northern boreal species. The vegetation at the Borden site is dominated by red maple (Acer rubrum), eastern white pine (Pinus strobus), large-tooth and trembling aspen (Populus grandidentata and Populus tremuloides) and white ash (Fraxinus americana [Lee et al., 1999; Teklemariam et al., 2009]. The fetch of largely uninterrupted forest extends to distances of 1.5–4 km in the southeastern and southwestern quadrants, and to 1 km in the northeastern direction. The northwestern fetch contains a white pine plantation and was not included in the footprint of eddy-covariance (EC) flux calculations at the site [Froelich et al., 2015]. The soil type is sandy loam [Gonsamo et al., 2015]. From 2008 to 2013, the average temperature was 8.64°C, precipitation was 808 mm/year and average daytime solar irradiance was 301 W/m².

3.3.2 Field measurements

Four dominant C3 broadleaf species (red maple, trembling aspen, large-tooth aspen and white ash) were sampled during the growing seasons of 2013, 2014 and 2015. Foliar chlorophyll was measured every 7 to 15 days from five leaf samples per species taken from top-of-canopy branches accessed directly from the 44 m flux tower. Branches were tagged to ensure repeatable sampling throughout the season, and the leaf samples were kept in cool, dark conditions during transport back to a laboratory for analysis. Leaf chlorophyll was extracted using N,N-dimethylformamide, and was analysed using a Shimadzu UV-1700 Spectrophotometer, with extinction coefficients specified by Wellburn (1994).

Leaf gas exchange measurements were made on the same days as chlorophyll sampling, using a LI-6400XT portable infrared gas analyser (LI-COR Inc., Lincoln, NE, USA). CO₂ response curves (A-Ci curve) of the leaves on the tagged branches were measured under an artificial saturated light source of 1800 µmol photons m⁻² s⁻¹ and stepwise ambient CO₂ concentrations of 400, 200, 100, 50, 400, 400, 600, 800, 1000, 1200, 1500, 1800 ppmv. \( \text{V}_{\text{max},0}^{25} \) and \( \text{J}_{\text{max},0}^{25} \), the \( \text{V}_{\text{max}}^{25} \) and \( \text{J}_{\text{max}}^{25} \) of the leaves at the top of the canopy, were calculated from the A-Ci curves using a curve-fitting tool developed by Kevin Tu.
(www.landflux.org) following Ethier and Livingston (2004) and then normalised to 25 °C [Sharkey et al., 2007].

Canopy structural parameters were also collected at 10 m intervals on the same days along a 100 m transect, extending from the flux tower in a North-South orientation. Effective LAI (L_e) values were obtained using the LAI-2000 plant canopy analyser (Li-Cor, Lincoln, NE, USA), while the element clumping index (Ω_E) was measured using the TRAC (Tracing Radiation and Architecture of Canopies) instrument [Chen and Cihlar, 1995]. The true LAI time sequences were then calculated as follows [Chen et al., 1999]:

\[
\text{LAI} = \frac{[(1 - \alpha) L_e \gamma_E]}{\Omega_E}
\]  

(3-1)

where α is the ratio of woody area to total area and γ_E is the ratio of needle area to shoot area. The ratio of woody area to total area (α = 0.17) (Croft et al, 2015a) accounts for the interception of radiation by branches and tree trunks that results in artificially high LAI values, and Ω_E is set at 0.95 based on the TRAC measurements. For broadleaf species, individual leaves are considered the foliage elements and γ_E is set at 1.

### 3.3.3 Flux and meteorological measurements

The eddy covariance measurements were made at the height of 33 m using a sonic anemometer (K-Type, Applied Technologies Inc., USA) coupled with a closed-path infrared gas analyzer (IRGA, model LI-6262, LI-COR Biosciences, USA) which is located in a temperature-controlled hut at the base of the tower. Both the IRGA and anemometer are operated at 10 Hz. High frequency eddy covariance fluxes were processed and aggregated into half-hourly fluxes using the method described in Froelich et al. (2015).

The half-hourly net ecosystem exchange (NEE, μmol m⁻² s⁻¹) is calculated as \( -(F_c + S_c) \), where \( F_c \) is the covariance between vertical wind velocity (w’) and atmospheric CO₂ concentration (c’). \( S_c \) is the rate of change of CO₂ storage per unit ground area in the air layer below the EC sensors. \( F_c \) is calculated from the WPL(Webb-Pearman-Leuning)-corrected vertical turbulent transport of the CO₂ mole fraction as \( F_c = \rho_a w' c'_{\text{WPL}} / M_a \), where the WPL accounts for density effects due to water vapor fluctuations [Webb et al., 1980], \( \rho_a \) is the density of air and \( M_a \) is the molecular weight of dry air. The storage
change $S_c$ is estimated as $S_c = \int_{z_0}^{z_{ec}} \Delta CO_2 \frac{\Delta t}{\Delta z} dz$, where $Z_{ec}$ is the height of the EC sensors at 33m and $t$ refers to time. Other terms such as horizontal advection are assumed to be negligible. NEE data were then filtered during periods of low turbulence [Staebler and Fitzjarrald, 2004] and when winds were from the direction of the short fetch (i.e. wind direction was $>285^\circ$ or $<20^\circ$) [Froelich et al., 2015]. A change point detection method (Barr et al., 2013) is used to derive the friction velocity threshold ($u^*$) to identify periods suitable for the application of the eddy covariance method to calculate fluxes. Gaps in the NEE data are filled using the method of Barr et al. (2004). During nights or cold periods respiration (RE) is simply computed as $RE = NEE$, while during warm periods or in cases of data gaps, RE is estimated using an empirical model based on air and soil temperature (Froelich et al., 2015). Half-hourly GPP is then calculated as $GPP = NEE + RE$; gaps in GPP are filled with an empirical model based on PAR (Barr et al., 2004; Froelich et al., 2015).

In conjunction with the CO$_2$ flux measurements, the EC system also measures the latent heat flux (LE, W m$^{-2}$) as $LE = L \rho_a w'q_{WPL}'$, where $L$ is the latent heat of vaporization (kJ g$^{-1}$), $\rho_a$ is the density of air (kg m$^{-3}$) and $w'q_{WPL}'$ is the WPL-corrected covariance between the vertical wind velocity and atmospheric water vapor mixing ratios (m s$^{-1}$ mol H$_2$O mol$^{-1}$ air). (Froelich et al., 2015; Teklemariam et al., 2009). Gaps in the LE measurements were not filled.

Several half-hourly auxiliary microclimate variables were also collected at the site to initialize and force the TBM. Wind speed (m s$^{-1}$) and wind direction were measured by the sonic anemometer mounted at the height of 33 m. Continuous air temperature ($^\circ$C) and relative humidity (%) were measured at various heights, but the 33 meter data were selected as model inputs. Incoming solar irradiance on the canopy was measured at the top of the flux tower [Froelich et al., 2015]. Soil moisture (m$^3$ m$^{-3}$) measured at the depth of 5 cm to 30 cm was used. Precipitation data (mm h$^{-1}$) were obtained from the nearest Environment Canada weather station that has hourly rainfall records (ID: Borden AWO), which is approximately 5.2 km away.
3.3.4 Terrestrial biophysical model

The Boreal Ecosystem Productivity Simulator (BEPS) is a two-leaf enzyme-kinetic model initially developed to estimate the carbon uptake and the water cycle of boreal ecosystems [Chen et al., 1999; Liu et al., 2003]. It has been substantially upgraded since its original release to simulate carbon and water fluxes at an hourly time step over various PFTs [Ju et al., 2006; Chen et al., 2007]. Several inter-model comparisons and site-level validations have shown that BEPS can produce reliable GPP and ET estimates [Amthor et al., 2001; Potter et al., 2001; Grant et al., 2006; Gonsamo et al., 2013]. The modules related to the utilization of Chl_leaf in BEPS are described in the following subsections.

3.3.4.1 Separation of sunlit and shaded leaves.

The two-leaf scheme is an abstraction of the instantaneous radiation regime in canopies. Leaves are separated into sunlit leaves and shaded leaves based on illumination and canopy geometry. Sunlit leaves tend to be light saturated because they receive both direct and diffuse solar radiation, while shaded leaves only receive diffuse radiation (Appendix A). The values of LAI_{sunlit} and LAI_{shaded} are calculated following the stratification scheme of Norman (1982) and Chen et al. (1999):

\[
\text{LAI}_{\text{sunlit}} = 2 \cos \theta (1 - \exp(-0.5 \Omega \text{LAI}_{\text{tot}} / \cos \theta)) \quad (3-2)
\]

\[
\text{LAI}_{\text{shaded}} = \text{LAI}_{\text{tot}} - \text{LAI}_{\text{sunlit}} \quad (3-3)
\]

where \( \theta \) is the solar zenith angle, LAI_{tot} is the total leaf area index of the canopy, and \( \Omega \) is the clumping index of this site, where \( \Omega = \Omega_E / \gamma_E \).

3.3.4.2 Derivation of \( V_{\text{max}}^{25} \) and \( J_{\text{max}}^{25} \) for sunlit and shaded leaves

In BEPS, \( V_{\text{max},0}^{25} \) is the input parameter to quantify the leaf photosynthetic capacity. The \( V_{\text{max}}^{25} \) values for sunlit and shaded leaves are each derived respectively based on the \( V_{\text{max},0}^{25} \) value, the vertical nitrogen profile of the canopy, the fraction of sunlit and shaded leaves
Leaf nitrogen content per unit leaf area $N(L)$ (g m$^{-2}$) generally decreases exponentially from the top to the bottom in a canopy following the long-term radiation distribution in the canopy (Equation 4):

$$N(L) = N_0 e^{-k_n L}$$

(3-4)

where the extinction coefficient $k_n = 0.3$ [De Pury and Farquhar, 1997], $N_0$ is the nitrogen content of leaves at the top-of-canopy, and $L$ is the canopy depth expressed as LAI from the top to a given height. $V_{max}^{25}$ is proportional to the leaf nitrogen content, and therefore it can be expressed as:

$$V_{max}^{25}(L) = V_{max,0}^{25} \chi_n N(L)$$

(3-5)

where $V_{max,0}^{25}$ is the $V_{max}^{25}$ of leaves at the top-of-canopy, and $\chi_n$ (m$^2$ g$^{-1}$) quantifies the relative change of $V_{max}^{25}$ to the leaf nitrogen content in the canopy, while $N(L)$ has the unit of g/m$^2$. The value of $\chi_n$, the mean value of $N_{area}$, its standard deviation, and the standard deviation of $V_{max}^{25}$ are provided for each PFT (Chen et al., 2012; Kattge et al., 2009); we have only used the values for deciduous broadleaf forests in this study. Since we do not have records specifically for $N_0$, $N_0$ is regarded as the sum of the mean $N_{area}$ and its standard deviation, which are provided by Kattge et al. (2009). Using Equation 4 and 5 and our estimated $N_0$, we can quantify the vertical patterns of $V_{max}^{25}$ in the canopy.

The fractions of sunlit ($f_{sun}$) and shaded leaves ($f_{sh}$) in the canopy change with the canopy depth are given by

$$f_{sun}(L) = \Omega e^{-kL}$$

(3-6)

$$f_{sh}(L) = 1 - \Omega e^{-kL}$$

(3-7)

respectively, where $k = G(\theta) \Omega / \cos \theta$. The projection coefficient of the canopy, $G(\theta)$, is taken as 0.5 assuming a spherical leaf angle distribution. We assume that $V_{max}^{25}$ of a representative sunlit or shaded leaf is equal to the mean $V_{max}^{25}$ value of the sunlit or shaded
leaf group. Therefore, the $V_{\text{max}}^{25}$ of a representative sunlit or shaded leaf is expressed through the following integrations:

$$V_{\text{max, sunlit}}^{25} = \frac{\int_0^L V_{\text{max, 0}} \chi_n N(L) f_{\text{sun}}(L) dL}{\int_0^L f_{\text{sun}}(L) dL} = V_{\text{max, 0}} \chi_n N_0 \frac{\int_0^L e^{-k_n L} e^{-k t} dL}{\int_0^L e^{-k t} dL}$$ (3-8)

$$V_{\text{max, shaded}}^{25} = \frac{\int_0^L V_{\text{max, 0}} \chi_n N(L) f_{\text{sh}}(L) dL}{\int_0^L f_{\text{sh}}(L) dL} = V_{\text{max, 0}} \chi_n N_0 \frac{\int_0^L e^{-k_n L} (1 - e^{-k L}) dL}{\int_0^L (1 - e^{-k L}) dL}$$

Several methods for estimating $V_{\text{max, 0}}^{25}$ are described in Section 3.3.5. $J_{\text{max}}^{25}$ for sunlit and shaded leaves are subsequently obtained using the $V_{\text{max}}^{25}$ built on our in-situ measurements (see Section 3.3.5).

3.3.4.3 Simulation of GPP and ET

BEPS calculates the instantaneous GPP as the sum of the photosynthesis from the sunlit leaves and the shaded leaves (Equation 10). Within each sunlit and shaded leaf group, all leaves are assumed to have the same physiological features; photosynthesis of a leaf group is then predicted using one representative leaf multiplied by LAI.

$$\text{GPP} = A_{\text{sunlit}} \times \text{LAI}_{\text{sunlit}} + A_{\text{shaded}} \times \text{LAI}_{\text{shaded}}$$ (3-10)

where $A_{\text{sunlit}}$ and $A_{\text{shaded}}$ are the instantaneous photosynthetic rates of a representative sunlit leaf and a representative shaded leaf, respectively. Following the Farquhar biochemical model, the instantaneous leaf photosynthetic rate in BEPS (Chen et al., 1999) is described as:

$$A = \min(W_c, W_f) - R_d$$ (3-11)

where $W_c$ is the rate of gross photosynthesis ($\mu$mol m$^{-2}$s$^{-1}$) when RuBP
carboxylase/oxygenase is saturated, and $W_j$ is the photosynthetic rate when RuBP-regeneration is limited by electron transport. $R_d$ is the dark respiration rate of CO$_2$ exchange.

\[
W_C = V_{\text{max}} \frac{C_i - \Gamma}{C_i + K} \quad (3-12)
\]

\[
W_j = J \frac{C_i - \Gamma}{4C_i + 8K} \quad (3-13)
\]

\[
R_d = 0.015V_{\text{max}} \quad (3-14)
\]

where $V_{\text{max}}$ is the maximum carboxylation rate ($\mu$mol m$^{-2}$s$^{-1}$) acquired from the prescribed $V_{\text{max}}^{25}$ and a temperature dependent function (Sharkey et al., 2007), $J$ is the electron transport rate ($\mu$mol m$^{-2}$s$^{-1}$). $C_i$ is the intercellular CO$_2$ concentration ($\mu$mol mol$^{-1}$), and $\Gamma$ is the CO$_2$ compensation point in the absence of dark respiration ($\mu$mol mol$^{-1}$). $K$ is a function of Rubisco enzyme kinetics described as $K_c/(1 + O_i/K_o)$, where $K_c$ and $K_o$ are Michaelis–Menten constants for CO$_2$ ($\mu$mol mol$^{-1}$) and O$_2$ (mmol mol$^{-1}$) and $O_i$ is the intercellular oxygen concentration (mmol mol$^{-1}$). The electron transport rate, $J$, is a function of the incoming photosynthetic photon flux density (PPFD in $\mu$mol m$^{-2}$s$^{-1}$) and the maximum electron transport ($J_{\text{max}}$):

\[
J = J_{\text{max}} \frac{\text{PPFD}}{(\text{PPFD} + 2.1J_{\text{max}})} \quad (3-15)
\]

Following a revised Ball-Woodrow-Berry equation (Ju et al., 2006), leaf stomatal conductance is quantified as:

\[
g_s = f_w \left( \frac{mA \cdot \text{RH}}{C_s} \right) + g_0 \quad (3-16)
\]

where $m$ is the dimensionless Ball-Woodrow-Berry coefficient set as 8, RH is the relative humidity (%), $C_s$ is the CO$_2$ concentration on the leaf surface, $g_0$ is the minimum conductance during the night, and $A$ is the rate of photosynthesis ($\mu$mol m$^{-2}$s$^{-1}$) of the representative sunlit or shaded leaf. The added $f_w$ variable is noted as the soil water stress factor (Appendix B).
After obtaining $g_s$, leaf-level transpiration is calculated by employing the Penman-Monteith equation in BEPS (Chen et al., 2007):

$$ T = \frac{\Delta (R_n - G) + \rho c_p VPD g_v}{\Delta + \left(1 + \frac{g_v}{g_s}\right) \gamma} \times \frac{1}{\lambda} \quad (3-17) $$

where $\lambda$ is the latent heat of evaporation of water (J kg\(^{-1}\)), $R_n$ is the instantaneous net radiation on the leaf surface (W/m\(^2\)) (Appendix A), $G$ is the rate of heat storage of the leaf which is a minimum close to 0 (W/m\(^2\)), $\rho$ is the density of air (kg m\(^{-3}\)), $c_p$ is the specific heat of air (J kg\(^{-1}\)°C\(^{-1}\)), $VPD$ is the vapor pressure deficit on the surface of a leaf (kPa), $\gamma$ is the psychrometric constant (kPa °C\(^{-1}\)), $g_v$ is the leaf boundary-layer resistance to water vapor (m s\(^{-1}\)), and $\Delta$ is the derivative of saturated vapor pressure with respect to the air temperature (kPa °C\(^{-1}\)). The leaf-level transpiration (mm s\(^{-1}\)) is upscaled to the canopy-level transpiration following an analogous protocol to the GPP upscaling.

$$ T = T_{\text{sunlit}} \times \text{LAI}_{\text{sunlit}} + T_{\text{shaded}} \times \text{LAI}_{\text{shaded}} \quad (3-18) $$

where $T_{\text{sunlit}}$ is the transpiration from a sunlit leaf and $T_{\text{shaded}}$ is the transpiration from a shaded leaf.

In addition to transpiration, BEPS also simulates evaporation from soil and from wet leaves (Chen et al., 2007). These processes are not regulated by stomata and therefore are less likely to be affected by the improvement in $V_{\text{max,0}}^{25}$ parameterization.

### 3.3.5 Modelling treatments

Empirical equations were developed to relate Chl\(_{\text{leaf}}\) to $V_{\text{max,0}}^{25}$ and $J_{\text{max,0}}^{25}$, based on field measurements (Croft et al., 2017), and were incorporated into BEPS for the purpose of improving GPP and ET simulations. In order to evaluate the improvements associated with utilising Chl\(_{\text{leaf}}\) (Case 4), we used three commonly employed $V_{\text{max,0}}^{25}$ treatments as reference cases (Case 1,2 and 3).

**Case 1:** constant $V_{\text{max,0}}^{25}$. Groenendijk et al. (2011) assimilated eddy covariance
measurements from Fluxnet into an inverse model of photosynthesis and transpiration to derive site-specific \( V_{\text{max}}^{25} \) and PFT-specific \( V_{\text{max}}^{25} \). Several BEPS studies have successfully used this set of \( V_{\text{max}}^{25} \) values for site-level GPP and ET simulations (Gonsamo et al., 2013; Chen et al., 2016; Luo et al., 2017, in review).

In this case, we used the value of 62 \( \mu \text{mol m}^{-2} \text{s}^{-1} \) for temperate deciduous broadleaf forest from Groenendijk et al. (2011) as the constant \( V_{\text{max}}^{25}_{\text{ref}} \) at the Borden Forest site. This value is comparable to another commonly cited value of 57.7 \( \mu \text{mol m}^{-2} \text{s}^{-1} \) for deciduous broadleaf forests reported by Kattge et al. (2009).

**Case 2: LAI-based \( V_{\text{max}}^{25}_{\text{ref}} \).** Some studies have found that seasonal patterns of \( V_{\text{max}}^{25} \) follow the seasonal patterns of LAI. We used a scheme developed by Ryu et al. (2011) to calculate \( V_{\text{max}}^{25}_{\text{ref}} \) on any given day during the growing season:

\[
V_{\text{max}}^{25}_{\text{ref}} = a V_{\text{max, ref}}^{25} + (1 - a) V_{\text{max, ref}}^{25} \frac{L_c - L_{\text{min}}}{L_{\text{max}} - L_{\text{min}}} \tag{3-19}
\]

where \( L_{\text{max}}, L_{\text{min}} \) and \( L_c \) are maximum, minimum and current LAI values over the year. \( V_{\text{max, ref}}^{25} \) is the value of maximum \( V_{\text{max}}^{25} \) during the growing season, which is regarded as 62 \( \mu \text{mol m}^{-2} \text{s}^{-1} \). The empirical variable \( a \) is set at 0.3 as in Ryu et al. (2011). The ratio component \( \frac{L_c - L_{\text{min}}}{L_{\text{max}} - L_{\text{min}}} \) ranges between 0 and 1.

**Case 3: SLA-based \( V_{\text{max}}^{25}_{\text{ref}} \).** The specific leaf area (SLA) of leaves at the canopy top is used in the Community Land Model version 4.0 (CLM4) to parameterize \( V_{\text{max}}^{25}_{\text{ref}} \) [Thornton and Zimmermann, 2007]. In this case:

\[
V_{\text{max}}^{25}_{\text{ref}} = N_{\text{area}} F_{LNR} F_{NR} \alpha_{R25} f(N) \tag{3-20}
\]

where \( N_{\text{area}} \) is the area-based leaf nitrogen concentration (g N m\(^{-2}\)), \( F_{LNR} = 0.09 \) is the fraction of leaf nitrogen in Rubisco (g N in Rubisco g\(^{-1}\) N) for temperate broadleaf deciduous trees, \( F_{NR} = 7.16 \) is the mass ratio of total Rubisco molecular mass to nitrogen in Rubisco (g Rubisco g\(^{-1}\) N in Rubisco), and \( \alpha_{R25} = 60 \) is the specific activity of Rubisco (\( \mu \text{mol CO}_2 \text{ g}^{-1} \text{ Rubisco s}^{-1} \)). A scaling factor, \( f(N) = 0.64 \), represents the effects of N
limitation. $N_{area}$ is calculated from mass-based leaf N concentration and SLA.

$$N_{area} = \frac{1}{CN_LSLA} \quad (3.21)$$

where $CN_L = 25$ is the leaf carbon-to-nitrogen ratio (g C g$^{-1}$ N) for broadleaf trees and SLA is specific leaf area (m$^2$ g$^{-1}$ C).

**Case 4:** Chl$_{leaf}$-based $V_{\text{max,0}}^{25}$. A straightforward way to implement Chl$_{leaf}$ into TBMs is through a robust relationship between Chl$_{leaf}$ and the photosynthetic parameters (i.e. $V_{\text{max}}^{25}$ and $J_{\text{max}}^{25}$). Croft et al. (2017) found a significant linear relationship between Chl$_{leaf}$ and $V_{\text{max,0}}^{25}$ ($r^2=0.76$, $p<0.001$) for the four deciduous species at the Borden site with an intercept close to zero (Equation 22 and 23).

$$V_{\text{max,0}}^{25} = 1.3 \times \text{Chl}_{\text{leaf}} + 3.72 \quad (3.22)$$

$$J_{\text{max,0}}^{25} = 1.85 \times V_{\text{max,0}}^{25} + 6.87 \quad (3.23)$$

For leaves at the top of the canopy, an increase of 1 μg cm$^{-2}$ of Chl$_{leaf}$ corresponds to 1.3 μmol m$^{-2}$ s$^{-1}$ in $V_{\text{max,0}}^{25}$. $V_{\text{max,0}}^{25}$ and $J_{\text{max,0}}^{25}$ are significantly related ($r^2=0.91$, $p<0.001$) by an empirical equation across four species (Figure 3-1).
**Figure 3-1.** Scatter plots of the measured leaf-level physiological parameters of C3 broadleaf leaves at the top of canopy at the Borden site in 2014 and 2015. (a) $V_{\text{max}_0}^{25}$ versus Chl$_{\text{leaf}}$; (b) $J_{\text{max}_0}^{25}$ versus $V_{\text{max}_0}^{25}$. The uncertainty of Chl$_{\text{leaf}}$ is smaller than 1 $\mu$g cm$^{-2}$. Source: Croft et al., 2017.

While the ranges of Chl$_{\text{leaf}}$, $V_{\text{max}_0}^{25}$ and $J_{\text{max}_0}^{25}$ each differ among the four species, the linear equations in Figure 3-1 fit the data points consistently, implying that it is reasonable to use a single equation to link Chl$_{\text{leaf}}$ and $V_{\text{max}_0}^{25}$ for broadleaf species. These two equations were used to incorporate Chl$_{\text{leaf}}$ into BEPS in Case 4.

### 3.4. Results

#### 3.4.1 Leaf traits and canopy fluxes

Figure 3-2 shows the variations of Chl$_{\text{leaf}}$, LAI and SLA that are used to constrain $V_{\text{max}_0}^{25}$ in the four modelling cases (Figure 3-2a), along with simulated $V_{\text{max}}^{25}$ and canopy flux measurements (Figure 3-2b) across two growing seasons. Chl$_{\text{leaf}}$, LAI and SLA were measured every 7-15 days during the two growing seasons; values between the sampling dates were determined by linear interpolation. The Chl$_{\text{leaf}}$ and SLA values were calculated based on the weighted composition of the major tree species at the Borden Forest (Teklemariam et al., 2009).
Figure 3-2. Seasonal variations of (a) Chl\textsubscript{leaf}, leaf area index (LAI), specific leaf area (SLA), and; (b) simulated $V_{\text{max,0}}^{25}$ for each modelling scenario where: Case 1 (a constant $V_{\text{max,0}}^{25}$), Case 2 (LAI-based $V_{\text{max,0}}^{25}$), Case 3 (SLA-based $V_{\text{max,0}}^{25}$) and Case 4 (Chl\textsubscript{leaf}-based $V_{\text{max,0}}^{25}$), alongside GPP and ET tower measurements, during 2013 and 2014 at the Borden Forest field site.

In 2013 and 2014, Chl\textsubscript{leaf} accumulated slowly, beginning in early May, reached its peak in July, with values ranging from 0.6 to 51.4 μg cm\textsuperscript{-2}. In contrast, leaves became fully expanded in a shorter time-frame (over 2 weeks in early-mid May), giving a steeper gradient in the early growing season LAI values. SLA decreased promptly at the beginning of growing seasons and stayed relatively constant from May to October. At the end of the growing season, the breakdown of chlorophyll began to exceed production in late August,
while leaf fall did not begin until October (Croft et al., 2014).

Among the four cases, the simulated Chl\textsubscript{leaf}-based $V_{\text{max}_0}^{25}$ (Case 4) appears to most closely follow the seasonal variations in measured GPP and ET (Figure 3-2b). In the modelled scenarios, simulated $V_{\text{max}_0}^{25}$ ranges from: Case 2 - 30.1 to 60.5 $\mu$mol m\textsuperscript{-2} s\textsuperscript{-1}; Case 3 - 35 to 70.2 $\mu$mol m\textsuperscript{-2} s\textsuperscript{-1}; and Case 4 - 4.4 to 70.6 $\mu$mol m\textsuperscript{-2} s\textsuperscript{-1}. Values of $V_{\text{max}_0}^{25}$ from these modelling cases are used in BEPS to simulate GPP and ET over two growing seasons at the Borden Forest.

3.4.2 Validation of the GPP and ET estimates from different $V_{\text{max}_0}^{25}$ constraints

The simulated GPP and ET results are shown for the four different modelling cases, in order to investigate the effects of the different biophysical constraints on $V_{\text{max}_0}^{25}$. Figure 3-3 shows linear regressions between the hourly measured fluxes and the hourly estimated fluxes from BEPS.

**Figure 3-3.** Scatter plots of the predicted and measured fluxes during 2013-2014 at the
Borden site. (a) to (d) refer to the GPP validation in Case 1 (a constant $V_{\text{max,0}}^{25}$), Case 2 (LAI-based $V_{\text{max,0}}^{25}$), Case 3 (SLA-based $V_{\text{max,0}}^{25}$) and Case 4 (Chlleaf-based $V_{\text{max,0}}^{25}$), respectively; (e) to (h) refer to the ET validation in Case 1, Case 2, Case 3 and Case 4, respectively.

Modelled results from Case 4 (Chlleaf-based $V_{\text{max,0}}^{25}$) show the strongest correlation ($r^2 = 0.91$) with measured GPP values, and the lowest RMSE of 0.13 g m$^{-2}$ h$^{-1}$ (Figure 3-3). The other three modelled scenarios reveal overestimated results at lower GPP values, which likely correspond to the start and the end of the growing season. Additionally, Case 1 (a constant $V_{\text{max,0}}^{25}$) shows underestimations at higher measured GPP values. The total simulated annual GPP values are 2123 g m$^{-2}$ y$^{-1}$, 1967 g m$^{-2}$ y$^{-1}$, 2124 g m$^{-2}$ y$^{-1}$ and 1729 g m$^{-2}$ y$^{-1}$ for Case 1 (a constant $V_{\text{max,0}}^{25}$), Case 2 (LAI-based $V_{\text{max,0}}^{25}$), Case 3 (SLA-based $V_{\text{max,0}}^{25}$) and Case 4 (Chlleaf-based $V_{\text{max,0}}^{25}$), respectively, while the annual GPP obtained from flux-tower measurements is 1719 g m$^{-2}$ y$^{-1}$. The bias of the estimated annual GPP for Case 4 is only 0.6% of the observed annual GPP, while the biases for Cases 1, 2 and 3 are 23.5%, 14.4% and 23.6%, respectively, of the observed annual GPP.

The Chlleaf-constrained $V_{\text{max,0}}^{25}$ also produces the strongest relationship between estimated and measured ET ($r^2 = 0.74$; RMSE = 0.07 mm h$^{-1}$). As with the GPP results, the estimated annual ET for Case 4 of 362 mm y$^{-1}$ is closest to the measured value of 370 mm y$^{-1}$, while the other annual ET results are 414 mm y$^{-1}$ (Case 1), 397 mm y$^{-1}$ (Case 2), 409 mm y$^{-1}$ (Case 3). The bias of the estimated annual ET from Case 4 is -2.2% of measured annual ET, while to the biases from Case 1, 2 and 3 are 11.9%, 7.3% and 10.5% of the annual observed ET. Overall, Case 4 is the most accurate method for simulating both the total amount and temporal variation of carbon and water fluxes.

3.4.3 Seasonality of the improvements in estimated GPP and ET

In order to investigate the seasonal variability in the improvements made to GPP and ET
estimations using Case 4 relative to Cases 1-3 (Figure 3-3), the monthly mean absolute errors (MAEs) and the relative MAEs of the modelled fluxes are shown for the four cases (Figure 3-4).

![Graphs showing monthly relative MAEs and MAEs between estimated and measured daily fluxes in 2013-2014, under the four modelling cases.](image)

**Figure 3-4.** The monthly relative MAEs (a, b) and the MAEs (c, d) between the estimated and measured daily fluxes in 2013-2014, under the four modelling cases. (a, c) GPP and (b, d) ET. The relative MAEs are the division of MAEs by daily flux values.

Figures 3-4a and 3-4b show that Case 4 introduces the smallest errors in most months for both GPP and ET estimates. Other cases display larger MAEs and relative MAEs than
Case 4, especially at the beginning and end of the growing seasons. In Figures 3-4c and 3-4d, the average MAEs of estimated GPP in spring (from April to June) of the four cases are 3.2 g m$^{-2}$ d$^{-1}$, 2.2 g m$^{-2}$ d$^{-1}$, 3.0 g m$^{-2}$ d$^{-1}$ and 1.8 g m$^{-2}$ d$^{-1}$ respectively; in autumn (from September to November), the average MAEs of estimated GPP for the four cases are 2.4 g m$^{-2}$ d$^{-1}$, 2.1 g m$^{-2}$ d$^{-1}$, 2.8 g m$^{-2}$ d$^{-1}$ and 1.8 g m$^{-2}$ d$^{-1}$, respectively. In the middle of the growing seasons, the improvements are not pronounced since the $V_{\text{max}}^{25}$ values derived from all cases are close to each other, and large errors remain. The major improvements in modelled ET also appear in spring (from April to June), where the MAEs are reduced from 0.8 mm d$^{-1}$ in Case 1, 0.7 mm d$^{-1}$ in Case 2 and 0.8 mm d$^{-1}$ in Case 3 to 0.6 mm d$^{-1}$ in Case 4. In fall (from September to November), small improvements on the order of 0.01 mm d$^{-1}$ in MAEs are observed from Case 1, 2 and 3 to Case 4. From June to August, the ET estimates show little differences among the four cases.

### 3.4.4 Assessing the relative contribution of Chl$_{\text{leaf}}$ to modelled GPP and ET

According to the modelling results, Chl$_{\text{leaf}}$-$V_{\text{max}}^{25}$ markedly improves the estimation of GPP and ET during the transitional periods of the growing season. However, the improvement in ET is not as great as the improvement in GPP (Figure 3-4). This decoupling effect of carbon and water fluxes suggests that there are differences in the sensitivities of GPP and ET to the forcing variables. In Figure 3-5, we explore the sensitivities of GPP and ET to five forcing variables of BEPS: incoming solar radiation ($R_{s}$), air temperature ($T_{\text{air}}$), relative humidity (RH), LAI and Chl$_{\text{leaf}}$. To assess the sensitivity of GPP and ET to certain variables, each variable is changed by -20%, -10%, 10% and 20% stepwise to drive BEPS (Case 4), while all other variables remain unchanged. The resulting changes in simulated GPP and ET are analyzed.
Figure 3-5. Sensitivity of GPP and ET to the incoming solar radiation ($R_s$), air temperature ($T_{air}$), relative humidity (RH), LAI and Chl$_{leaf}$ at the Borden site. (a) GPP; (b) ET.

Figure 3-5a shows that GPP is most sensitive to the changes in Chl$_{leaf}$. A 20% variation in Chl$_{leaf}$ leads to a change in annual GPP as large as 200 g m$^{-2}$ y$^{-1}$. In addition to Chl$_{leaf}$, $R_s$ and LAI are linearly linked to GPP with similar order of importance. There is an optimal $T_{air}$ range for GPP, with higher or lower $T_{air}$ damping the carbon assimilation rate. RH has the least influence on GPP. In contrast, ET is most sensitive to the climatic variables, $T_{air}$ and RH, while variables used to describe the plant functional status (i.e. LAI and Chl$_{leaf}$) drive smaller changes in ET (Figure 3-5b). As such, the variation in ET resulting from the changing Chl$_{leaf}$ and LAI in spring and autumn is overshadowed by the variation from changing temperature and humidity. Consequently, ET is less improved than GPP by incorporating Chl$_{leaf}$ based $V_{m}^{25}$ in the TBM.

3.4.5 The impact of light environment within the canopy on leaf physiology

This study incorporates Chl$_{leaf}$ into a two-leaf model to consider the complex light environment in forests and the accompanying physiological traits of sunlit and shaded
leaves. Light intensity within the canopy affects the relative allocation of nitrogen between light-harvesting chlorophyll molecules and Rubisco in the Calvin cycle. Consequently, it is important to consider how this dynamic partitioning of nitrogen affects the integration of chlorophyll into the two-leaf BEPS model, and how this partitioning impacts carbon and water exchange. Figure 3-6 shows the contrasting modelled solar irradiance for sunlit and shaded leaves over the two growing seasons.

**Figure 3-6.** Seasonal patterns of the average daily incoming solar radiation (W/m²) on sunlit and shaded leaves simulated by BEPS in 2013 and 2014.

The daytime average solar radiation on sunlit leaves ranges from 105 to 316 W m² from DOY 100 to DOY 300, with an average value of 216 W m². In contrast, the solar irradiance on shaded leaves ranges from 22 to 75 W m² in the same period, with an average of 36 W m². This 6-fold difference in the light environment results in a different role of chlorophyll between the sunlit and shaded leaves, as shown in Figure 3-7.
Figure 3-7. BEPS modelled leaf photosynthesis and transpiration rates and their relationship with Chl<sub>leaf</sub>. Sunlit (open dots) and shaded leaves (solid dots) are compared. Ten-day averages of the modelled photosynthesis and transpiration rates in the growing seasons are used.

For modelled sunlit leaves, an increase of 1 µg cm<sup>-2</sup> in Chl<sub>leaf</sub> increases the rate of photosynthesis by 0.22 µmol m<sup>-2</sup> s<sup>-1</sup> and the rate of transpiration by 0.04 mmol m<sup>-2</sup> s<sup>-1</sup> in sunlit leaves. In contrast, photosynthesis and transpiration of shaded leaves are less sensitive to changes in Chl<sub>leaf</sub>: a 1 µg cm<sup>-2</sup> increase in Chl<sub>leaf</sub> leads to only a 0.07 µmol m<sup>-2</sup> s<sup>-1</sup> increase in photosynthesis and a 0.01 mmol m<sup>-2</sup> s<sup>-1</sup> increase in transpiration. Importantly, Figure 3-7 suggests that chlorophyll plays a lesser role in determining GPP and ET in shaded leaves than in sunlit leaves. In line with the different role of chlorophyll, sunlit and shaded leaves shows differences in other physiological traits (Figure 3-8).
Figure 3-8. Physiological differences, from BEPS simulations, between sunlit and shaded leaves. (a) $V_{\text{max}}^{25}$ versus leaf chlorophyll content; (b) leaf-level photosynthesis versus transpiration; (c) the seasonal variation in LUE.

The seasonal average $V_{\text{max}}^{25}$ of a sunlit leaf is 34% higher than that of a shaded leaf, with the biggest difference of 53% occurring in the middle of the growing seasons (Figure 3-8a). According to our modelling results, the carbon and water fluxes of leaves are strongly coupled, and the water use efficiency (WUE) of the shaded leaf group is similar to that of sunlit leaves (Figure 3-8b). For a unit area of leaves, 1 mmol of transpired water could incur 5.5 and 5.8 µmols of carbon assimilation for sunlit leaves and shaded leaves, respectively. However, light use efficiency (LUE) of leaves fluctuates with Chl leaf [Houborg et al., 2011; Croft et al., 2015]. According to modelled results from BEPS, the
LUE of sunlit leaves ranges from 0.02 to 0.62 g C m\(^{-2}\) MJ\(^{-1}\) APAR with an average of 0.35 g m\(^{-2}\) MJ\(^{-1}\). Shaded leaves are more efficient in photosynthesis with a LUE ranging from 0.02 to 1.54 g C m\(^{-2}\) MJ\(^{-1}\), averaging 1.04 g C m\(^{-2}\) MJ\(^{-1}\) (Figure 3-8c). These LUE values are within the range of observed and modelled LUE reported previously reported in the literature [Medlyn, 1998; Yuan et al., 2007]. Our finding -- that the modelled LUE of shaded leaves is almost 3 times the modelled LUE of sunlit leaves -- echoes the theory that plants are more capable of utilising diffuse radiation than direct radiation [Gu et al., 2002], since shaded leaves only receive diffuse solar radiation for photosynthesis. Another explanation is that leaves tend to decrease their LUE with increasing intercepted solar radiation. While sunlit leaves usually receive more radiation than shaded leaves, the LUE of sunlit leaves is expected to be lower than that of shaded leaves.

### 3.5. Discussion

**3.5.1 Improved seasonal representation for GPP and ET simulations**

The comparison of the four modelling cases demonstrates that Chl\(_{\text{leaf}}\) is the best constraint for \(V_{\text{max}}^{25}\) in the TBM. Seasonal changes in \(V_{\text{max}}^{25}\) represent changes in the leaf Rubisco content, which is believed to synchronize with the leaf nitrogen pool [Evans, 1989a]. In previous studies, leaf structural parameters such as LAI and SLA were used in models to represent seasonal changes in \(V_{\text{max}}^{25}\), assuming that LAI and SLA can provide robust representations of the leaf total nitrogen content. However, the seasonal change in leaf total nitrogen varies from the trend of the leaf photosynthetic nitrogen pool [Croft et al., 2017; Migita et al., 2007; Warren et al., 2000], which usually accounts for 50% of the total leaf nitrogen, due to disparity between leaf physiological and morphological development shown in Figure 3-4 [Croft et al., 2014b]. As a consequence, it is important to consider leaf photosynthetic nitrogen, rather than leaf total nitrogen, and its relationship to Rubisco content. In this study, we find that Chl\(_{\text{leaf}}\) can serve as a direct proxy of the leaf photosynthetic nitrogen pool at the canopy top, and thus, can constrain the seasonal variation of \(V_{\text{max}}^{25}\). The results also suggest that the accumulation and degradation rates for the photosynthetic nitrogen pool (as represented by Chl\(_{\text{leaf}}\)) differ from the rates of changes
of leaf and canopy structural components (i.e. LAI and SLA). Consequently, the seasonality of GPP and ET is better explained by Chl$_{leaf}$ than by SLA or LAI, especially during these transitional periods when the disparity between leaf structure and leaf physiology reaches its maximum. In the middle of the growing season, the simulations from four cases are quite similar to each other because leaves are structurally and physiologically mature.

The seasonality of GPP and ET is also influenced by climatic factors. Though carbon and water fluxes are regulated by the stomata [e.g. Baldocchi 1994; Leuning 1990], our sensitivity analysis shows that GPP and ET are not tightly coupled as expected because of their different sensitivities to non-biological factors. Figure 3-5 demonstrates that Chl$_{leaf}$ is the dominant factor in controlling the carbon uptake, whereas air temperature and atmospheric mositure demand place the strongest controls on water exchange. This finding was also noted in previous sensitivity studies on several TBMs, which also found that V$_{max}^{25}$ is a more influential factor for carbon flux modelling than for water flux modelling [Alton et al., 2007; Ryu et al., 2011]. Wang and Dickinson [2012] reviewed observational ET studies across the globe and suggested that available energy and vapor pressure deficit are the two most important factors in determining ET in temperate and boreal forest ecosystems (such as at the boreal Borden site). Consequently, it is reasonable that the ET simulations are improved less than the GPP simulations after incorporating Chl$_{leaf}$-based V$_{max}^{25}$ in the TBM.

3.5.2 Physiological differences between sunlit and shaded leaves

Many studies have compared the physiological traits of sun- and shade-adapted leaves [e.g. Anderson et al., 1988; Boardman, 1977; Chow and Anderson, 1987] and reported that the constituents of photosynthetic components vary with leaf light environment. Under a saturated light environment, leaves tend to assign more nitrogen to produce Rubisco needed in dark reactions of photosynthesis, while shaded leaves will invest more nutrients in the production of the light-harvesting apparatus to capture the photons needed in light reactions. In a canopy, nitrogen is distributed proportionally from top leaves to bottom
leaves following the long-term radiation gradient. The sunlit part of the crown tends to have higher $N_{\text{area}}$ to produce more photosynthetic components (e.g. Rubisco) in comparison to the shaded part of the canopy. As such, the leaf physiological traits acclimate to leaf light environments to maximize the overall productivity of the whole canopy. Understanding the influence of the light environment on leaves is important before upscaling modelled GPP and ET from leaf to canopy in structurally complex forests.

Estimations from BEPS capture the physiological difference between sunlit and shaded leaves (Figures 3-7 and 3-8). With an increase in Chl$_{\text{leaf}}$, $V_{\text{max}}^{25}$ and photosynthesis by shaded leaves increase more slowly than those of sunlit leaves, because shaded leaves are predominantly located at the bottom of the canopy where the relatively low nitrogen content constrains the content of Rubisco. Therefore, the linear relationship between Chl$_{\text{leaf}}$ and $V_{\text{max}}^{25}$ has a smaller slope for shaded leaves than for sunlit leaves. This finding echoes previous experiments reporting changes in Chl$_{\text{leaf}}$-$V_{\text{max}}^{25}$ relationships under different light environments (Table 3-1).

Table 3-1. Slopes in the linear Chl$_{\text{leaf}}$-$V_{\text{max}}^{25}$ relationships reported in the literatures.

Ratios of $V_{\text{max}}^{25}$ / Chl$_{\text{leaf}}$ in some studies are regarded as slopes. Molecular mass of Chl$_{\text{leaf}}$ is 893.5 g mol$^{-1}$.

<table>
<thead>
<tr>
<th>Species</th>
<th>Solar irradiance on leaves (μmol m$^{-2}$ s$^{-1}$)</th>
<th>Slopes in Chl$<em>{\text{leaf}}$-$V</em>{\text{max}}^{25}$ relationships (μmol m$^{-2}$ s$^{-1}$ $V_{\text{max}}^{25}$ per μg cm$^{-2}$ Chl$_{\text{leaf}}$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cucumis</td>
<td>1000</td>
<td>1.82</td>
<td>[Evans, 1989a]$^a$</td>
</tr>
<tr>
<td></td>
<td>550</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>Phaseolus</td>
<td>820</td>
<td>2.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>220</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>1.32</td>
<td></td>
</tr>
<tr>
<td>Maize</td>
<td>1.12</td>
<td></td>
<td>[Houborg et al., 2015]</td>
</tr>
<tr>
<td>Soybean</td>
<td>1.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorghum</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat</td>
<td>2.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Chl&lt;sub&gt;leaf&lt;/sub&gt;–&lt;tex&gt;V_{max}^{25}&lt;/tex&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common bean</td>
<td>1.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotton</td>
<td>1.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>1.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barley</td>
<td>3.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco</td>
<td>2.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deciduous broadleaf tree species</td>
<td>1265&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sunlit leaves</td>
<td>972&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shaded leaves</td>
<td>162&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*<sup>a</sup> J<sub>max</sub> is converted into <tex>V_{max}^{25}</tex> by simply dividing 2;  
*<sup>b</sup> Measured daytime solar irradiance for leaves on top of the canopy;  
*<sup>c</sup> The average daytime solar irradiance on leaves, modelled from BEPS, during the period of the growing seasons.

The experimental studies included in Table 3-1 demonstrate that leaf light environments affect the slopes of Chl<sub>leaf</sub> –<tex>V_{max}^{25}</tex> relationships to a degree that is almost comparable to the influence brought by species type. While an increase of 1 μg cm<sup>-2</sup> in Chl<sub>leaf</sub> may result in 1 to 3 μmol m<sup>-2</sup> s<sup>-1</sup> of increases in <tex>V_{max}^{25}</tex> for different species, a change in the light environment for a given species could dampen its Chl<sub>leaf</sub> –<tex>V_{max}^{25}</tex> slope by a factor between 21% and 44%. Meanwhile, our modelled results show that shaded leaves have a 34% lower Chl<sub>leaf</sub> –<tex>V_{max}^{25}</tex> slope than sunlit leaves for deciduous broadleaf trees, which is similar to the results from previous limited experimental studies [Evans, 1989b]. Therefore, we suggest that our two-leaf TBM is capable of describing the variations in light environments and is reliable for incorporating Chl<sub>leaf</sub> in structurally complex forests.

### 3.6. Conclusions

Chlorophyll molecules embedded in the light harvesting apparatus are responsible for absorbing solar energy for photosynthesis, and they are also related, through sharing of the leaf nitrogen pool, to the enzyme Rubisco and the cytochrome b6f complex (Cty f) needed by photosynthesis. Based on the key role of chlorophyll in photosynthesis, Chl<sub>leaf</sub> is regarded as a valuable proxy of the photosynthetic parameters – <tex>V_{max}^{25}</tex> and J<sub>max</sub> – used in TBMs. However, since sun- and shade- adapted leaves have different strategies in the nitrogen allocation to components of the photosynthetic apparatuses (i.e. chlorophyll and
Rubisco), understanding the influence of the light environment on leaves is important before upscaling modelled GPP and ET from leaf to canopy in structurally complex forests. In this study, we incorporate a time series of measured Chl$_{leaf}$ into a two-leaf TBM named BEPS. The following conclusions are drawn:

1. Chl$_{leaf}$ provides a reliable constraint on the seasonal variations of $V_{\text{max}}^{25}$ and $J_{\text{max}}^{25}$ at a forest site. By incorporating Chl$_{leaf}$-based $V_{\text{max}}^{25}$ into BEPS, the biases of simulated annual GPP and annual ET are considerably reduced during spring and fall and the temporal correlations between simulated and measured fluxes are considerably improved relative to three cases using constant $V_{\text{max}}^{25}$, using LAI-based $V_{\text{max}}^{25}$ and using SLA-based $V_{\text{max}}^{25}$. The major improvements in GPP estimates are witnessed in spring and fall, when MAEs are reduced from between 2.2-3.2 to 1.8 g C m$^{-2}$ d$^{-1}$ and from between 2.1-2.8 to 1.8 g C m$^{-2}$ d$^{-1}$, respectively. MAEs of modelled ET also shrink from 0.7-0.8 to 0.6 mm d$^{-1}$ in spring while simulated ET does not improve very much in autumn.

2. The two-leaf TBM with $V_{\text{max}}^{25}$ constrained by Chl$_{leaf}$ is capable of capturing the physiological differences between sunlit leaves and shaded leaves in a forest stand. According to modelled results, a 1 $\mu$g cm$^{-2}$ increase in Chl$_{leaf}$ corresponds to a 1.2 $\mu$mols m$^{-2}$ s$^{-1}$ increase in $V_{\text{max}}^{25}$ in sunlit leaves and a 0.77 $\mu$mols m$^{-2}$ s$^{-1}$ increase in $V_{\text{max}}^{25}$ in shaded leaves. The result is in line with the plant physiological studies that found that plants optimize the nitrogen allocations to different photosynthetic components according to the light environment, in order to maximize the canopy-scale CO$_2$ assimilation rate.

This study demonstrates, for the first time, the importance of using chlorophyll in TBMs to reduce the uncertainties in carbon and water flux estimates for forested ecosystems. It also incorporates Chl$_{leaf}$ within a two-leaf scheme to account for the complex light environments inside forest canopies and quantifies the physiological difference between sunlit and shaded leaves.

Acknowledgement
This study is financially supported by a Discovery Grant and a Strategic Grant from the Natural Science and Engineering Council of Canada. We are grateful to Prof. Sarah Finkelstein from the University of Toronto for providing the spectrophotometer to help with chlorophyll measurements.

3.7 Appendix

3.7.1 Net radiation on sunlit and shaded leaves

In every hourly time step of BEPS, the whole canopy was divided into four groups of leaves based on the location and radiation features of the leaves, namely sunlit leaves in the overstorey, shaded leaves in the overstory, sunlit leaves in the understorey and shaded leaves in the understorey [Chen et al., 1999; Liu et al., 2003]. The leaves in each group have identical features so BEPS could use one leaf to represent one group. Net radiation on a leaf is composed by three sources:

\[ R_{n,i} = R_{dir,i} + R_{dif,i} + R_{l,i} \]  \hspace{1cm} (3-A1)

where \( R_n \) is the total net radiation on a certain leaf, \( R_{dir} \), \( R_{dif} \) and \( R_l \) refers to the net direct incoming solar radiation, net diffuse solar radiation and net longwave radiation on this leaf. \( i \) refers to one of the four types of leaves. For a shaded leaf, \( R_{dir} = 0 \).

In order to differentiate the incoming solar radiation into a direct and diffuse part, a semi-empirical equation is applied:

\[ \frac{s_{dif}}{s_g} = \begin{cases} 0.943 + 0.734r - 4.9r^2 + 1.796r^3 + 2.058r^4 & r < 0.8 \\ 0.13 & r \geq 0.8 \end{cases} \]  \hspace{1cm} (3-A2)

\[ S_{dir} = S_g - S_{dif} \]  \hspace{1cm} (3-A3)

where \( S_g \), \( S_{dir} \) and \( S_{dif} \) are global solar radiation, incoming direct solar radiation and diffuse solar radiation. \( r \) is a parameter used to quantify the cloudiness in sky

\[ r = \frac{s_g}{s_0 \cos \theta} \]  \hspace{1cm} (3-A4)
\( S_0 \) is the solar constant set as 1362 \( \text{W/m}^2 \), \( \theta \) is the solar zenith angle in this hourly time step.

The net direct shortwave radiation on the sunlit representative leaf in the overstorey or understorey of the canopy is:

\[
R_{\text{dir, o, sunlit}} = R_{\text{dir, u, sunlit}} = (1 - \alpha_L)S_{\text{dir}}\cos\alpha / \cos \theta
\]  

(3-A5)

where \( \alpha_L \) is the albedo of leaves. \( \alpha \) is the mean leaf-sun angle which is fixed at 60° when the canopy has a spherical leaf distribution.

On the other hand, the net diffuse shortwave radiation on the four groups of the leaves are approximated respectively as:

\[
R_{\text{dif, o, sunlit}} = R_{\text{dif, o, shaded}} = (1 - \alpha_L)(S_{\text{dif}} [1 - e^{-0.5 \Omega LAI_o / \cos \tilde{\theta}_o}] / LAI_o + C_o)
\]  

(3-A6)

\[
R_{\text{dif, u, sunlit}} = R_{\text{dif, u, shaded}} = (1 - \alpha_L)(S_{\text{dif}} e^{-0.5 \Omega LAI_u / \cos \tilde{\theta}_u} [1 - e^{-0.5 \Omega LAI_u / \cos \tilde{\theta}_u}] / LAI_u + C_u)
\]  

(3-A7)

\( LAI_o \) and \( LAI_u \) denote the LAI value of the overstorey and the understorey, \( C_o \) and \( C_u \) are used to quantify the multiple scattering of the direct solar radiation from the leaf (Chen et al., 1999)

\[
C_o = 0.07 \Omega S_{\text{dir}} (1.1 - 0.1 LAI) e^{-\cos \theta}
\]  

(3-A8)

\[
C_u = 0.07 \Omega S_{\text{dir}} e^{-0.5 \Omega LAI_o / \cos \tilde{\theta}_o} (1.1 - 0.1 LAI_u) e^{-\cos \theta}
\]  

(3-A9)

\( \tilde{\theta}_o \) and \( \tilde{\theta}_o \) are the representative zenith angles for diffuse radiation transmission of the overstorey and understorey leaves and slightly dependent on the corresponding LAI [Liu et al., 2003]:

\[
\cos \tilde{\theta} = 0.537 + 0.025 LAI
\]  

(3-A10)

The net longwave radiation on these leaves is calculated as:
\[ R_{L,o,sunlit} = R_{L,o,shaded} = \frac{1}{L_{AI}} \left\{ \left[ \epsilon_o \left[ \epsilon_a \sigma T_a^4 + \epsilon_u \sigma T_u^4 \left( 1 - e^{-0.5LAI_o \epsilon_o \cos \theta_o} \right) \right] + \epsilon_g \sigma T_g^4 e^{-0.5LAI_o \epsilon_o \cos \theta_o} \right] - 2 \epsilon_o \sigma T_o^4 \left( 1 - e^{-0.5LAI_o \epsilon_o \cos \theta_o} \right) + \epsilon_o (1 - \epsilon_u) (1 - e^{-0.5LAI_o \epsilon_o \cos \theta_o}) \right\} \]

\[ R_{L,u,sunlit} = R_{L,u,shaded} = \frac{1}{L_{AI}} \left\{ \left[ \epsilon_u \left[ \epsilon_a \sigma T_a^4 e^{-0.5LAI_o \epsilon_o \cos \theta_o} + \epsilon_o \sigma T_o^4 \left( 1 - e^{-0.5LAI_o \epsilon_o \cos \theta_o} \right) \right] + \epsilon_u (1 - \epsilon_o) \left[ \epsilon_u \sigma T_u^4 \left( 1 - e^{-0.5LAI_o \epsilon_o \cos \theta_o} \right) + \epsilon_u \sigma T_u^4 \left( 1 - e^{-0.5LAI_o \epsilon_o \cos \theta_o} \right) \right] \right\} \]

where \( \sigma \) is the Stephen-Boltzmann constant equals to \( 5.67 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4} \). \( \epsilon_a, \epsilon_o, \epsilon_u \) and \( \epsilon_g \) are the emissivity of the atmosphere, overstory, understory and ground surface, respectively. \( \epsilon_o, \epsilon_u \) and \( \epsilon_g \) are prescribed as 0.98, 0.98 and 0.95 according to [Chen and Zhang, 1989; Chen et al., 1989], and \( \epsilon_a \) is computed as \( \epsilon_a = 1.24 \left( \frac{e_a}{T_a} \right)^{1/7} \) [Brutsaert, 1982], where \( e_a \) and \( T_a \) are water vapor pressure in mbar and temperature of the atmosphere in K. \( T_o, T_u \) and \( T_g \) are the temperatures of the overstory, the understory and ground in K.

### 3.7.2 Quantification of the soil water stress factor

To assess the effect of soil water deficit on stomatal conductance \( (fw) \), a scaling factor based on the ratio of the real time available water in soil to the maximum plant available water [Wigmastoa et al., 1994; Wang and Leuning, 1998; Chen et al., 2005] is used to implement the soil water information.

\[ fw = \begin{cases} 0 & \theta_{sw}(z) - \theta_{wp} < 0 \\ \frac{\theta_{wp} - \theta_{sw}(z)}{\theta_f - \theta_{wp}} & 0 \leq \theta_{sw}(z) \leq \theta_f \\ \frac{1}{1} & \theta_{sw}(z) > \theta_f \end{cases} \]  

(3-B1)
where $\theta_{sw}(z)$ is the soil water content of layer $z$, $z$ often refers to the top 30 cm based on the availability of the soil water measurements. $\theta_{wp}$ and $\theta_{fc}$ are the wilting point and the field capacity ($m^3/m^3$) of the soil layer. $\theta_{wp}$ and $\theta_{fc}$ are derived by the soil texture information provided on the Fluxnet archive (http://fluxnet.ornl.gov/), the patterns of multi-year soil moisture measurements and the algorithm developed by Saxton and Rawls (2006).

3.8 References


Chen, J. M., B. J. Yang, and R. H. Zhang (1989), Soil thermal emissivity as affected by its water content and surface treatment, *Soil science (USA)*.


Groenendijk, M. et al. (2011), Seasonal variation of photosynthetic model parameters and


Porcar-Castell, A., E. Tyystjärvi, J. Atherton, C. Van Der Tol, J. Flexas, E. E. Pfundel, J. Moreno, C. Frankenberg, and J. A. Berry (2014), Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges,


Ryu, Y. et al. (2011), Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales, *Global Biogeochemical Cycles, 25*(4), n/a-n/a, doi:10.1029/2011GB004053.


doi:10.5589/m07-037.
Chapter 4
Chlorophyll dynamics explains spatial and temporal variabilities in terrestrial carbon uptake and evapotranspiration

Preamble: this chapter is close to be submitted. The authors of this paper are: X. Luo, H. Croft, J. M. Chen and L. He. XL collected the flux data, performed the data analysis and wrote the paper, XL, HC and JC collectively developed the idea, HC provided the algorithm to derive leaf chlorophyll content. HC and JC also contributed to writing, LH provided the global clumping index data.

4.1 Abstract
Terrestrial carbon and water fluxes are driven by a range of abiotic and biotic factors. State-of-the-art terrestrial biosphere models (TBMs) use synoptic descriptions of biogeochemical processes to represent the impacts of these factors on terrestrial fluxes, such as gross primary productivity (GPP) and evapotranspiration (ET). Whilst TBMs have captured the mean values of global GPP and ET relatively well, the spatial and temporal variabilities in GPP and ET are not well accounted for, mainly due to the underrepresentation of biotic factors in TBMs. Leaf maximum carboxylation rate (Vcmax25) is a key biotic factor prescribed in TBMs to determine CO₂ assimilation rates and leaf stomatal conductivity for water transport, but the paucity of its measurements has long plagued the simulation of fluxes. This study uses leaf chlorophyll content (LCC) derived from remotely-sensed data to account for spatial and temporal variations in Vcmax25 within a TBM framework. Results from the TBM with and without LCC are validated against measurements from 124 eddy-covariance towers (554 site-years) of FLUXNET. Using LCC in a TBM reduces the biases of estimated GPP and ET in 61% and 59% of the site-years, respectively, with especially large improvements for biomes with strong seasonal cycles (e.g. deciduous forest, croplands and grasslands). In addition to the Vcmax25 adjustment imposed by LCC seasonal patterns, the spatial variability of
LCC acts as an equally important part in reducing the errors of estimated fluxes by capturing the spatial variations of Vcmax25, especially during the summer. This study presents the first case of integrating satellite-derived LCC into a TBM at the global scale. Our results demonstrate the critical role of LCC in describing the variabilities in the terrestrial carbon uptake and ET and the necessity of including LCC in future TBMs.

4.2 Introduction

The accurate description of biosphere-atmosphere CO₂ and H₂O exchange is central to predicting the responses of the land surface to future climate [IPCC, 2013; Keenan et al., 2013]. Currently, the total carbon assimilated by the biosphere via photosynthesis (gross primary productivity; GPP) is reported to be around 120-130 petagrams per year [Beer et al., 2010; Zhang et al., 2016], whilst terrestrial evapotranspiration (ET) transports 60 to 70 × 10^3 km^3/yr of water to the atmosphere [Oki and Kanae, 2006; Jung et al., 2010; Mueller et al., 2013; Jiang and Ryu, 2016]. Although the mean values of annual GPP and ET are constrained to a relatively small range, the spatial and temporal variability in GPP [Keenan et al., 2012a; Baldocchi et al., 2016] and ET [Wang and Dickinson, 2012] are less certain due to a diverse range of abiotic factors (i.e. climatic anomalies, soil moisture stress and fire) and biotic factors (i.e. plant phenology, nutrient availability, insects and ecosystem alternation) controlling them.

Terrestrial biosphere models (TBMs) are the principal tools for providing regional and global estimates of terrestrial GPP and ET [Friend et al., 2007; Beer et al., 2010], using the numerical representations of the biotic and abiotic factors to drive the physical and physiological processes on the land surface in a synoptic and accurate manner. The leaf maximum carboxylation rate (normalized to 25 degrees; Vcmax25) is a key biotic factor considered in TBMs. Vcmax25 corresponds to the activity of the photosynthetic enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) in the Calvin-Benson cycle, and in turn dominates the net carbon assimilation rate of leaves (A_n) [Farquhar et al., 1980]. Since A_n controls leaf stomatal conductance (g_s) [Ball, 1988; Leuning, 1990], water fluxes from leaves are subsequently controlled by Vcmax25. Vcmax25 is arguably the most critical biotic parameter required by TBMs [Bonan et al., 2011; Schaefer et al., 2012].
Vcmax25 has been shown to vary according to seasonality and environmental stresses, and between and within PFTs [Xu and Baldocchi, 2003; Groenendijk et al., 2011a; Keenan et al., 2012b; Croft et al., 2017]. However, the direct quantification of Vcmax25 requires time-consuming measurement of A-Ci curves at the leaf level [Xu and Baldocchi, 2003], precluding the availability of large spatially- and temporally-distributed datasets needed for global modelling. Many TBMs have used fixed Vcmax25 values for a given PFT [Wullschleger, 1993; Ryu et al., 2011] as a compromise, or from derived relationships with an apparent measurable variable. Kattge et al. [2009] compiled global A_n and Vcmax25 records in the TRY plant trait database and use leaf nitrogen content (N_area) as a proxy for Vcmax25. Other studies have explored the possibility of assimilating flux measurements into inverse TBMs to model optimal Vcmax25 for individual eddy-covariance sites [Groenendijk et al., 2011b; Zheng et al., 2017]. Using these methods, however, Vcmax25 is still assigned to PFTs or a limited number of sites with lack of response to seasonality and environmental stresses.

Chlorophyll is a crucial component of plant photosynthesis, through harvesting photons and the producing biochemical energy to support the Calvin-Benson cycle [Porcar-Castell et al., 2014; Alton, 2017]. Total leaf nitrogen content is partitioned between structural and photosynthetic fractions, and leaf chlorophyll content (LCC), defined as the total chlorophyll a and chlorophyll b content per one half of the total leaf area, can act as a proxy of the status of photosynthetic nitrogen pool shared by Rubisco [Evans, 1989; Hikosaka and Terashima, 1996; Croft et al., 2017]. Croft et al. [2017] recently found a single relationship between LCC and Vcmax25 of four deciduous tree species in a temperate forest. Earlier studies from Houborg et al. [2013] and Houborg et al. [2015] also developed some linear relationships between Vcmax25 and LCC for C3 and C4 cropland species, using N_area as an intermediate. The integration of these LCC-Vcmax25 relationships into TBMs were found to reduce errors and improve GPP and ET simulations at some individual cropland sites [Houborg et al., 2013] and forest site [Luo et al., in review]. At the global scale, a reliable way to map Vcmax25 may be through mapping LCC rather than N because LCC is a pigment with strong absorption characteristics in
visible bands, while radiation absorption by nitrogen in leaves is confounded with many other factors and only non-structural nitrogen components.

An essential operational factor for integrating LCC, as a proxy for $V_{cmax25}$, into global-scale modelling efforts is the accuracy of LCC values from remotely sensed data. Many studies have found that LCC is sensitive to changes in reflectance in the red-edge spectral region, which results from an abrupt gradient from red to near-infrared bands [Gitelson et al., 1996]. Empirical spectral vegetation indices have been used to quantify local-scale variations in LCC within the same PFT [Gitelson, 2005; Wu et al., 2008; Croft et al., 2014]. However, in addition to LCC, satellite-derived empirical indices also contain reflectance contributions from leaf area index (LAI), clumping, leaf angle distribution, non-photosynthetic material and understory vegetation [Haboudane et al., 2002; Simic et al., 2011]. To remove the effects of canopy structure in LCC derivations, a physically-based modelling approach has been demonstrated to accurately model LCC for a range of different PFTS [Zhang et al., 2008; Croft et al., 2013; Houborg et al., 2015a].

In this study, we use a physically-based modelling approach to derive LCC from MEdition Resolution Imaging Spectrometer (MERIS) data. The LCC series provides a basis to depict the variability of $V_{cmax25}$ with a 7-day interval and 300 m resolution. Modelled GPP and ET from the TBM using LCC constrained $V_{cmax25}$ and fixed $V_{cmax25}$ values according to PFT was compared against measured GPP and ET data at 124 eddy-covariance sites from FLUXNET for 554 site-years [Papale et al., 2006; Baldocchi, 2008], covering nine different PFTs. This study will evaluate the role of LCC in modelling biosphere-atmosphere fluxes across a comprehensive range of PFTs and identify the uncertainties involved in using LCC within TBMs.

4.3 Materials and method

4.3.1 Flux and meteorological dataset

FLUXNET2015 is the newest effort to refine the flux measurements from eddy-covariance towers of different regional networks, and publishes the data in a standard format. It provides gap-filled GPP and ET records and their concurrent meteorological records for
over 200 sites. In this study, we use the FLUXNET2015 Tier 1 dataset released in November 2016, which contains 146 sites across nine major PFTs: cropland (CRO), deciduous broadleaf forest (DBF), evergreen broadleaf forest (EBF), evergreen needleleaf forest (ENF), mixed forest (MF), grassland (GRA), savanna and woody savanna (SAV), closed and open shrubland (SH), and wetland (WET). Depending on the availability of scenes from MERIS that are used for the LCC derivation, 124 sites (554 site-years) are solicited for further modelling and validation.

Half-hourly and hourly meteorological records in the FLUXNET2015 dataset are used to drive a TBM—the Boreal ecosystem productivity simulator (BEPS). Gap-filled solar radiation (SW_IN_F), air temperature (TA_F), vapor pressure deficit (VPD_F), precipitation (P_F) and wind speed (WS_F) are selected as the forcing variables for BEPS, to make sure that the TBM can produce continuous hourly results. The FLUXNET2015 dataset uses two streams of methods, namely variable u* threshold for each year (VUT) and constant u* threshold across years (CUT), to partition Net Ecosystem Exchange (NEE) into Ecosystem Respiration (RE) and GPP. Many versions of the GPP records are created by using different percentiles of u* in each stream of method. Among those GPP versions, we regard GPP_NT_VUT_REF as the measured GPP, which is created by using the VUT method and the original u* values, and we only use nighttime data to parameterize a respiration model that is then applied to the whole dataset to estimate RE [Reichstein et al., 2005]. To get maximum amount of ET records for validation, the gap-filled latent heat (LE_F_MDS) is chosen and converted into the unit of mm/h.
Quality assessment (QA) flags and observational uncertainties are considered in the validation process. Only the half-hourly and hourly records labelled as good measurement (QA=0) and gap-filled with confidence (QA=1) are taken forward for validation. The high-quality hourly values are then summed to daily, weekly and annually time-scales. Consequently, the daily/weekly/annual GPP and ET we provided in this research could be different from the values provided in FLUXNET2015 dataset which aggregates values from all QA categories. The observational uncertainty of GPP is propagated from that of NEE, noted as NEE_VUT_REF_RANDUNC in the dataset; and the uncertainty of ET measurement is LE_RANDUNC, which is also converted into mm/h.

4.3.2 Remote sensing data

4.3.2.1 Leaf area index

Daily LAI series are required for individual sites to drive a TBM and derive LCC. We selected the Copernicus Global Land Service GEOV1 LAI product derived from SPOT-VEGETATION satellite, which has a global coverage of LAI from 1999 to the 2014, at
10-day temporal intervals and a spatial resolution of 1 km [Baret et al., 2013]. The GEOV1 LAI derivation algorithm is based on a neural-network practised by the CYCLOPES Version 3.1 and MODIS Collection 5 LAI products, in such manner that it capitalizes on the development and validation of already existing products. The GEOV1 LAI product considers clumping as a weighted contribution of CYCLOPES and MODIS products, where MODIS LAI accounts for clumping at plant and canopy scales [Knyazikhin et al., 1998], and the CYCLOPES algorithm accounts for landscape clumping by considering fractions of mixed pixels. One possible drawback of the GEOV1 product pertains to needle-leaf forests, where shoot clumping is not accounted for in the algorithm, but a recent validation study indicates that GEOV1 outperforms most existing products in accuracy, with good spatial consistency, smooth temporal profile and reliable magnitude for bare areas and dense forests [Camacho et al., 2013]. We use the Locally Adjusted Cubic-spline Capping (LACC) method [Chen et al., 2006] to interpolate and smooth the discontinuous LAI into daily LAI series. LACC is capable of identifying the points in a time series contaminated by clouds, sub-pixel clouds and other atmospheric effects, and replacing the contaminated points through temporal interpolation between adjacent valid points based on the seasonal trajectory. In a recent comparison with five other interpolation methods, including the Savitzky-Golay filter and the Fourier-based approach, LACC showed better performance in protecting key points and removing noise in time series of vegetation indices [Liu et al., 2017].

4.3.2.2. Clumping index

The clumping index ($\Omega$) describes the non-randomness of the leaf distribution in the canopy [Chen et al., 1997]. It is a key parameter to drive the canopy radiation modules in TBM and the LCC derivation algorithm. Site-specific CI values are obtained from a global foliage clumping index map produced from the MODIS BRDF products [He et al., 2012]. The map has a resolution of 500 m, which is comparable to the size of regular tower footprints.

4.3.2.3. MEedium Resolution Imaging Spectrometer (MERIS) satellite data
MEdium Resolution Imaging Spectrometer (MERIS) satellite surface reflectance data were selected for LCC derivation because of their inclusion of chlorophyll-sensitive red-edge bands, relatively fine spatial resolution (300 m), and good temporal resolution (7 days) [Rast et al., 1999]. The MERIS surface reflectance series are produced by a series of pre-processing steps, including radiometric, geometric and BRDF atmospheric correction, pixel identification, and atmospheric correction with aerosol retrieval. There are 13 bands (spectral resolution = ~10 nm) in the visible, red-edge and near infra-red bands sampled in the reflectance dataset. MERIS covers the complete years from 2003 to 2011. We extract the surface reflectance in all bands at every flux tower site as the input for the LCC derivation algorithm. In order to retain enough sampling points to detect the seasonal patterns of LCC, only the site-years that have more than 10 MERIS surface reflectance records are considered in this study.

4.3.3 Terrestrial biosphere model

The TBM used in this study is the Boreal Ecosystem Productivity Simulator (BEPS), which is a two-leaf enzyme kinetic model that has been intensively used to simulate carbon and water fluxes across different biomes [Wang et al., 2004; Gonsamo et al., 2013]. It has participated in several cross-model validation studies and been proved to be one of the better performing models for GPP and ET simulation [Grant et al., 2006; Schaefer et al., 2012]. Parameterizations of BEPS were demonstrated in detail in previous studies [Chen et al., 1999, 2007]. A recent study has upgraded BEPS to integrate LCC (a.k.a. BEPS-LCC) in modelling fluxes in a deciduous broadleaf forest [Luo et al., in review]. Figure 4-2 depicts the model structure of BEPS-REF and BEPS-LCC.
Figure 4-2. Flow chart of the structure of BEPS-REF based on leaf area index (LAI) and clumping index (CI) and BEPS-LCC based LAI, CI and leaf chlorophyll content (LCC).

Meteorological variables (i.e. solar radiation, temperature, VPD, precipitation and wind speed) are the inputs for BEPS. The solar radiation and leaf temperature are separately calculated for sunlit and shaded leaves through a two-leaf scheme using LAI and CI [Chen et al., 1999]. Leaf-level photosynthesis is then obtained from the embedded Farquhar’s model with the input of leaf traits. Here BEPS-REF uses seasonally constant Vcmax25 provided by Kattge et al. [2009] while BEPS-LCC uses LCC-based Vcmax25 to quantify photosynthesis ability. After using Farquhar’s model to calculate An, Stomatal conductance gs is in turn derived from the Ball-Woodrow-Berry equation and used in the Penman-Monteith model to get leaf transpiration. In feedback, the changes in leaf energy budget affect the leaf temperature and adjust the photosynthesis rate.

After the leaf-level calculation, canopy-scale photosynthesis and transpiration are acquired by multiplying the leaf-level fluxes by the corresponding sunlit and shaded LAI through the two-leaf scheme. Soil evaporation is an integrated part of ET simulation in BEPS, but it is not regulated by stomata and thus not sensitive to LCC. Soil texture is prescribed for each site using a global map (https://www.soilgrids.org/) to quantify the hydrological and thermal properties of soil and to simulate soil water content. We apply a linear soil stress
factor derived from multilayer soil water content to simulate the stomatal response to possible droughts [Luo et al., in review].

4.3.4 Derivation of LCC

A two-step process-based algorithm is used to derive LCC from MERIS surface reflectance. The first step is to model leaf spectra from satellite-derived canopy reflectance using inverse canopy reflectance models, and the modelled leaf reflectance spectra is then inputted into a leaf radiative transfer model to obtain LCC. Two different canopy reflectance models are selected according to the structural characteristics of the vegetation present. For spatially 'clumped' vegetation types (i.e. deciduous and coniferous trees, shrubs) we select the 4-Scale geometrical–optical model [Chen and Leblanc, 1997]. For homogenous canopies that can be treated as turbid media, we use the SAIL model [Verhoef, 1984]. Individual look up tables (LUT) were created for each plant functional type, in order to invert the canopy models in which the relationships between canopy reflectance and leaf reflectance are built and categorized by canopy structural variables. The LUT approach is selected to optimise computational resources and reduce problems associated with local minima, given sufficiently sampling of the variable space [Jacquemoud et al., 2009]. The second step is to derive leaf biochemical constituents from the modelled leaf reflectance using an inverse leaf radiative transfer model (PROSPECT) [Jacquemoud and Baret, 1990]. This two-step approach has been successfully demonstrated over various sites and biomes [Zhang et al., 2008; Simic et al., 2011; Croft et al., 2013, 2015a].

4.3.5 Derivation of Vcmax25 from LCC

Linear equations are used to translate the weekly LCC into weekly Vcmax25 (Table 4-1). The Vcmax25-LCC relationships used in DBF, ENF, MF, SAV, WET are based on the work of Croft et al. [2017], who built a single and significant linear relationship across four deciduous broadleaf tree species. The cropland algorithm is based on the approach by Houborg et al. [2015], who used a mechanistic framework to build an equation to link LCC and Vcmax25 via area-based leaf total nitrogen content (Narea), in which a single equation was developed for croplands. This relationship is also applied for another
herbaceous type (GRA) in this study. We then expand this mechanism to build \( \text{Vcmax25-LCC} \) equations for EBF and SH, using a comprehensive \( \text{Vcmax25-Narea} \) database compiled by Kattge et al. [2009] and a constant ratio between LCC and \( \text{Narea} \) provided by Evans [1989] (Table 4-1).

**Table 4-1.** Linear regression equations used to derive \( \text{Vcmax25} \) from LCC (\( \text{Vcmax25} = a \times \text{LCC} + b \))

<table>
<thead>
<tr>
<th>PFTs</th>
<th>( a ) (( \mu \text{mol}/\text{m}^2/\text{s} ) per ( \mu \text{g}/\text{cm}^2 ))</th>
<th>( b ) (( \mu \text{mol}/\text{m}^2/\text{s} ))</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBF, ENF, MF, SAV, WET</td>
<td>1.3</td>
<td>3.72</td>
<td>[Croft et al., 2017]</td>
</tr>
<tr>
<td>CRO, GRA</td>
<td>1.98</td>
<td>12.5</td>
<td>[Houborg et al., 2015b]</td>
</tr>
<tr>
<td>SH</td>
<td>0.95</td>
<td>14.71</td>
<td>[Kattge et al., 2009] &amp; [Evans, 1989]</td>
</tr>
<tr>
<td>EBF</td>
<td>0.66</td>
<td>6.99</td>
<td></td>
</tr>
</tbody>
</table>

### 4.4 Results

#### 4.4.1 Seasonality of LCC and \( \text{Vcmax25} \) according to PFT

Variations in the satellite-derived LCC, LAI and \( \text{Vcmax25} \) for the major PFTs are revealed in Figure 4-3. LCC is derived from MERIS satellite reflectance data (see Section 4.3.4), LAI is acquired from GEOV1 dataset, and the seasonality of \( \text{Vcmax25} \) values is calculated according to the relationship with LCC shown in Table 4-1.
Figure 4-3. Mean LCC, LAI and Vcmax25 for nine PFTs per week. The LCC, LAI and Vcmax25 variations at sites in the South Hemisphere is reversed to permit an comparison with the seasonal pattern at sites in the Northern Hemisphere.

All PFTs except EBF show different degrees of LCC seasonality, with EBF remaining high and relatively steady at circa 66 μg/cm² throughout the whole year. LCC values for the forests types DBF, MF and ENF vary from circa 20 μg/cm² to 60 μg/cm², with peak values usually occurring in summer. The peaks of LCC decrease in the order of DBF>MF>ENF, indicating differences in the photosynthetic capacities between broadleaf and needleleaf species. Whilst the curve of CRO LCC shows a strong phenological variation, CRO LCC maintains its values above 40 μg/cm² across the year because of multi-cropping systems. SH and SAV demonstrate shorter growing seasons where their LCC increases quickly in spring and begin declining shortly afterwards. LCC-based
Vcmax25 shows the same seasonal variability as LCC; however, because the relationships between Vcmax25 and LCC vary by PFT, the magnitudes of Vcmax25 vary in a different order than LCC. CRO and GRA, for example, demonstrate the highest Vcmax25 - close to 100 μmol/m²/s and 80 μmol/m²/s, respectively, while Vcmax25 of other PFTs fluctuates around 40-50 μmol/m²/s. The structural parameter LAI demonstrates different phenology to that of the physiological parameter LCC, suggesting the development of leaf morphology varies from the accumulation or degradation of pigments and enzymes in leaves.

To assess the robustness of the LCC-based Vcmax25, we compare them against the widely used Vcmax25 values acquired from the TRY database [Kattge et al., 2009] and the Vcmax25 values inverted from flux measurements [Groenendijk et al., 2011b] in Table 4-2.

**Table 4-2.** The mean and standard deviations of site-average Vcmax25 for each PFTs, unit is μmol/m²/s, SD refers to standard deviation.

<table>
<thead>
<tr>
<th>PFTs</th>
<th>LCC-based Vcmax25 [Kattge et al., 2009]</th>
<th>[Groenendijk et al., 2011b]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>CRO</td>
<td>95.4</td>
<td>41.1</td>
</tr>
<tr>
<td>DBF</td>
<td>50.8</td>
<td>23.5</td>
</tr>
<tr>
<td>EBF</td>
<td>50.6</td>
<td>9.8</td>
</tr>
<tr>
<td>ENF</td>
<td>45</td>
<td>22.3</td>
</tr>
<tr>
<td>MF</td>
<td>49.9</td>
<td>19.1</td>
</tr>
<tr>
<td>GRA</td>
<td>69.9</td>
<td>26.4</td>
</tr>
<tr>
<td>SAV</td>
<td>43.7</td>
<td>23.8</td>
</tr>
<tr>
<td>SH</td>
<td>55.8</td>
<td>22.2</td>
</tr>
<tr>
<td>WET</td>
<td>40.7</td>
<td>22.6</td>
</tr>
</tbody>
</table>

a use the value for nonoxisols tropical trees.
b weighted average of evergreen shrubs and deciduous shrubs
The Vcmax25 derived from LCC is constrained to a similar range as the values reported by the other two studies. GRA and CRO have the largest Vcmax25 among PFTs. According to Table 4-2, the mean LCC-based Vcmax25 values are distributed between 40 μmol/m²/s and 55 μmol/m²/s for all PFTs except for GRA and CRO. All forest species have very similar average Vcmax25 values across the year though their seasonal variabilities are different. The SD values of LCC-based Vcmax25 are also close to the values reported by other sources, which suggests that LCC-based Vcmax25 shows a reasonable spatial variability.

4.4.2 Examining the advantage of using LCC to model GPP and ET

GPP and ET estimated by BEPS-LCC and BEPS-REF for each site-year are validated against the measured fluxes. R² and bias are used as the indicators to evaluate the improvement of BEPS-LCC over BEPS-REF (Figure 4-4).
Figure 4-4. The reduction of bias and the improvement of R$^2$ for GPP (a, c) and ET (b, d) estimated by BEPS-LCC compared to the fluxes estimated by BEPS-LCC.

Out of the 554 site-years, 61% of site-years show reduced bias for GPP estimation and 59% for ET by using BEPS-LCC. In terms of GPP bias (Figure 4-4 a), five PFTs have more than 50% of the site-years showing decreased bias after using LCC, such as CRO (93%), DBF (66%), GRA (63%), SH (88%), WET (59%), while EBF (25%), ENF (47%), MF (24%) and SAV (24%) show the opposite. In terms of ET bias (Figure 4-4 b), the PFTs and their percentages of site-years showing improvements are: CRO (69%), DBF (54%), EBF (12%), ENF (76%), MF (23%), GRA (64%), SAV (17%), SH (89%), WET (55%).

Similar to the distribution of the bias reductions, 58% of site-years show improved R$^2$ for GPP estimation and 58% for ET estimation by using BEPS-LCC. The PFTs and their corresponding percentage of site-years showing improved R$^2$ for GPP estimations are (Figure 4-4 c) CRO (82%), DBF (85%), EBF (65%), ENF (23%), MF (26%), GRA (62%), SAV (94%), SH (46%), WET (84%); for ET estimations are (Figure 4-4 d) CRO (74%), DBF (64%), EBF (45%), ENF (39%), MF (61%), GRA (62%), SAV (76%), SH (50%), WET (70%).

The seasonal variations of the improvements by BEPS-LCC are evaluated according to the distribution of biases in Figures 4-5 and 4-6. We plot the weekly biases of estimated GPP and ET from BEPS using LCC-based Vcmax25 (BEPS-LCC) and BEPS using fixed Vcmax25 [Kattge et al., 2009](BEPS-REF) against EC-derived GPP and ET, across the year. The stand deviation (SD) of the biases and weekly uncertainties of GPP and ET observations are included in each panel.
Figure 4-5. The biases of estimated GPP along with the range of biases (defined as ±SD). Blue buffers refer to the observational uncertainty.

Major reductions in modelled GPP error for BEPS-LCC estimates are clearly visible for CRO, DBF, MF and GRA sites. In particular, BEPS-LCC corrects the overestimations of GPP for COR, GRA and SH in the middle of the growing seasons for some site-years, improves DBF GPP at the start and middle of the growing season. The GPP improvements in summer are likely to be driven by the down-regulation of peak Vcmax25 by LCC at some sites where the prescribed Vcmax25 was high, rather than driven by the seasonal variations of Vcmax25 added by LCC. It indicates that LCC can be a robust representation
of the spatial variations of \( V_{cmax25} \). WET and SH also show some improvement in reducing the biases. The biases of GPP estimations are mostly constrained by using LCC in BEPS. ENF is the only PFT shows no improvement, with larger mean biases appearing at the beginning of the seasons.

Figure 4-6. The biases of estimated ET along with the range of biases (defined as \( \pm SD \)). Blue buffers refer to the observational uncertainty.
Figure 4-6 demonstrates that the biases in ET estimations of some PTFs from BEPS-LCC are constrained to a lesser degree than those of GPP. CRO, GRA, MF and ENF show smaller biases in the middle of the seasons while DBF shows decreased biases at the beginning of the growing season by using LCC in BEPS. Other PFTs do not show statistically strong changes from BEPS-REF to BEPS-LCC.

Additionally, a comprehensive comparison of the modelling results against the FLUXNET2015 dataset is provided as Supporting Information (Section 4.7.1). BEPS-LCC shows stronger performance than BEPS in terms of acquiring higher $R^2$ and lower RMSEs, in either temporal or spatial validation. Uncertainties of the modelling results propagated from the LCC derivation do not affect the validation results (Section 4.7.2)

4.5 Discussion

4.5.1 Assessment of the Vcmax25-LCC relationships

The translation of LCC into Vcmax25 is a key step in integrating LCC into TBMs. While very few studies have examined the correlation between LCC and Vcmax, they have reported linear equations to describe the relationship [Houborg et al., 2013; Croft et al., 2017], assuming that the photosynthetic nitrogen pool is allocated proportionally to LCC and Rubisco. A non-linear relationship has also been proposed by Alton [2017], in which he assumed a linear relationship between LCC and Jmax25 (a.k.a. maximum electron transport rate) and a non-linear relationship between Vcmax25 and Jmax25. The latter assumption challenges existing studies which suggest a linear Jmax25-Vcmax25 relationship [e.g. Wullschleger, 1993]. According to Alton et al. [2017], the non-linear Jmax25-Vcmax25 equation varies 12% from its linear form, and this degree of uncertainty may propagate to the final Vcmax25-LCC equation, affecting simulations of fluxes.

This study uses the Vcmax25-LCC relationships reported by Croft et al. [2017] and Houborg et al. [2015] for majority of PFTs. For PFTs (i.e. EBF and SH) where direct Vcmax25-LCC relationships are not available, we combined the known Vcmax25-Narea relationship and the LCC-Narea ratio to derive PFT-specific Vcmax25-LCC relationship
In addition, we tried a third method by linking the site-specific LCC derived from satellite and the site-specific Vcmax25 derived from an inverse TBM [Groenendijk et al. 2011b]. The Vcmax25-LCC relationships acquired from the three methods are compiled in Table 4-3.

**Table 4-3. Vcmax25-LCC relationships (Vcmax25 = a*LCC + b) derived using different methods.**

<table>
<thead>
<tr>
<th>PFTs</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>CRO</td>
<td>1.98</td>
<td>12.5</td>
<td>2.43</td>
</tr>
<tr>
<td>DBF</td>
<td>1.3</td>
<td>3.72</td>
<td>1.22</td>
</tr>
<tr>
<td>EBF</td>
<td></td>
<td></td>
<td>0.66</td>
</tr>
<tr>
<td>ENF</td>
<td>1.3</td>
<td>3.72</td>
<td>0.75</td>
</tr>
<tr>
<td>MF</td>
<td>1.3</td>
<td>3.72</td>
<td></td>
</tr>
<tr>
<td>GRA</td>
<td>1.98</td>
<td>12.5</td>
<td>1.68</td>
</tr>
<tr>
<td>SAV</td>
<td>1.3</td>
<td>3.72</td>
<td></td>
</tr>
<tr>
<td>SH</td>
<td></td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>WET</td>
<td>1.3</td>
<td>3.72</td>
<td></td>
</tr>
</tbody>
</table>

Note: unit of a is μmol/m²/s per μg/cm², unit of b is μmol/m²/s.

Method 1 includes equations reported by literature [Croft et al., 2017; Houborg et al., 2015]; Method 2 uses the Vcmax25-Na relationships introduced in Kattge et al. [2009] and the LCC-Na ratio introduced by Evans [1989] to build the Vcmax25-LCC equations; Method 3 links the site-specific satellite-derived LCC in this study and site-specific Vcmax25 from Groenendijk et al. [2011b] to build the equations.

All three methods show that CRO and GRA have the highest efficiency in using LCC for carboxylation (measured by Vcmax25), suggesting that herbaceous species may impose a different nitrogen allocation strategy than others. The difference of the slopes for forest types between the three methods leads to an important source of uncertainty in the Vcmax25-LCC relationships – leaf light environments. Leaves inside the canopy can be generally classified into two types – sunlit and shaded leaves [Sinclair et al., 1976; Norman, 1982]. Sunlit leaves are light saturated where the photosynthesis rate is limited by the abundance of enzyme Rubisco for the dark reactions of the Calvin-Benson cycle. In contrast, shaded leaves require more photons, and its photosynthesis is limited by the
photons harvested by chlorophyll molecules. Plants tend to optimize their nitrogen allocation between different photosynthetic components to maximize their total canopy CO₂ assimilation rate [e.g. Field, 1983; Kull, 2002]. Thus, sunlit leaves allocate more nitrogen to Rubisco and less nitrogen to chlorophyll, whilst more nitrogen will be assigned to the production of LCC in shaded leaves. Given the light acclimation of canopies, the Vcmax25-LCC relationship of shaded leaves show a lower slope than that of sunlit leaves [Luo et al., in review].

The single relationship reported by Croft et al. [2017] is based on the measurement of leaves on top of the canopy, which are all sunlit leaves located in light-saturated conditions. However, in Method 2, since the Vcmax25, Aₙ and Nₘ area data are compiled from various sources, it is hard to guarantee that they are all collected from leaves under the same light environments. Since the conifer forests are always more clumped than broadleaf forests [Chen et al., 1997], there are more shaded leaves inside the conifer canopies, making the leaf Vcmax25-LCC slope lower on average. Method 3 uses flux measurements in an inverse TBM to derive Vcmax25, but it applies a big-leaf scheme which is unable to separate sunlit and shaded leaves. While big-leaf TBMs are found to underestimate GPP [Chen et al., 1999; Sprintsin et al., 2012], an inverse TBM will tend to overestimate Vcmax25 to compensate for errors in its modelling structure. Further study in this direction would require Vcmax25 derived from an inversion of some two-leaf TBMs [Zheng et al., 2017].

4.5.2 Uncertainty of LCC for evergreen PFTs

Whilst the satellite-derived LCC of EBF, ENF and MF appears to represent the seasonal LCC phenologies for these PFTs, the addition of LCC in BEPS fails to improve the GPP and ET simulation for these PFTs. It implies that LCC of evergreen trees may relate to Vcmax25 in a different way than that of the deciduous species.

Deciduous species experience a full growth cycle on an annual basis, generating and dismantling photosynthetic components (i.e. LCC and Rubisco) in every cycle. The
variability of their photosynthetic ability is directly linked to the amount of LCC and Rubisco in chloroplasts. Evergreen species, on the other hand, will adjust LCC and Rubisco content between the leaves of different ages. Katahata et al. [2007] reported that needle leaves belonging to ages of 1, 2, and 3 years show different seasonal trends in LCC and Vcmax25. Young needles show an increasing trend in LCC while the old needles show the opposite. Warren and Adams [2001] found that old needles contain less Rubisco but almost the same N than young needles by studying 6 different age groups of maritime pines. By studying Douglas fir, Ethier et al. [2006] suggested that the reduction in photosynthetic capacity of old leaves is not triggered by the decreased allocation of N towards Rubisco, but caused by the decreased activation state of Rubisco and a proportional down-regulation of electron transport. As for EBF, a recent study in Amazonia demonstrated that leaf age composition explains 27% of the variation in photosynthesis, and that young leaves have higher light use efficiency than the old leaves [Wu et al., 2016]. These studies restrict us from applying one single LCC-Vcmax25 relationship to ENF and EBF without the consideration of the age structure of leaves in the canopy. Satellite-derived LCC quantifies the chlorophyll content of leaves on top of the canopy. It is a mixed signal of LCC of different age groups for evergreen forests, because the leaf radiative transfer model PROSPECT (which used to derive LCC) does not consider leaf aging effects when deriving LCC. The aging effect presents more challenges for expanding application of LCC in evergreen species.

Evergreen species may also be able to keep their maximum canopy photosynthetic ability relatively stable because they can dynamically optimize Vcmax25 and LCC inside the canopy according to leaf age. Their seasonal variation in the photosynthesis abilities is more likely to be adjusted by radiation and temperature. For example, GPP simulated for ENF from BEPS-LCC shows larger negative biases at the beginning of the seasons than the simulated GPP from BEPS (Figure 4-5). This indicates that temperature adjustment is enough to constrain the seasonal variations of GPP without using varying Vcmax25. Previous modelling results based on 22 TBM s also demonstrate the best performance at conifer forests in temperate regions [Schwalm et al., 2010], implying less adjustments are needed for Vcmax25 of ENF.
4.5.3 Relative sensitivities of GPP and ET fluxes to biophysical and environmental drivers

Carbon and water fluxes to and from the biosphere are collectively regulated by biotic factors (i.e. LCC, LAI) and abiotic factors (i.e. radiation, temperature, humidity and wind). However, different sensitivities of water and carbon fluxes to LCC may lead to the weak improvement for ET relative to GPP after using LCC in the model. Figure 4-9 explores the daily variations explained by variables through simple linear regressions.

**Figure 4-7.** Daily variations of (a) GPP and (b) ET explained by variables including LCC, LAI, incoming solar radiation (Rs), air temperature (Tair), relative humidity (RH), precipitation(P) and wind speed (Ws) based on measurements from 124 flux towers.

According to Figure 4-9(a), LCC explains 13%-49% of the daily variations of GPP across nine PFTs. GPP of CRO, DBF, SAV and WET shows higher correlations to LCC than those of ENF, EBF and MF, because CRO, DBF, SAV and WET are commonly recognized to have strong phenological cycles. LAI, incoming solar radiation and temperature are the other major factors that are highly correlated to GPP variations. Since
LAI essentially reflects the vegetation structure in the same way as NDVI [Croft et al., 2015b], the great R² between GPP and LAI echoes the fact that some semi-empirical GPP models using vegetation indices such as NDVI and EVI as proxies of GPP can work reasonably well on some flux tower sites [e.g. Xiao et al., 2004; Wu et al., 2010].

Figure 4-9(b) shows that LCC explains 15%–44% of the variation in ET depending on PFTs, while DBF and SAV show the strongest power of LCC in explaining ET variability. Whilst LAI has the strongest influence on GPP, the dominant factor controlling ET is incoming solar radiation. Models such as the Priestley and Taylor equation [Priestley and Taylor, 1972] have been using solar radiation as the sole input to calculate ET and it has been proved useful across several PFTs [Vogel et al., 1995; Fisher et al., 2008]. The atmospheric vapour demand, indicated by RH, also demonstrates more influence on ET than it does on GPP. Sensitivity analysis on some TBMs has implied that GPP and ET are not as tightly coupled as expected because of the difference in their response to abiotic factors [Alton et al., 2007; Ryu et al., 2011]. Even though CO₂ and H₂O both use stomata as the exchange passage, the carbon flux can actively regulate the stomata openness while water fluxes passively escape through stomata driven by radiation and humidity. We also find the seasonal patterns of GPP and ET is more sensitive to LAI than to LCC, indicating the amount (area) of leaves is a first order factor in determining the dynamic of fluxes while the status of leaves plays a secondary role.

4.6 Conclusion

This study presents the first efforts to use satellite-derived LCC in a TBM at the global scale, in order to better represent Vcmax25 variations within TBMs and to improve GPP and ET estimation. Through the validation of modelled results with data from 124 FLUXNET sites, the following conclusions are drawn:

1. Satellite-derived LCC produces strong seasonal and spatial dynamics of Vcmax25 across various PFTs and FLUXNET sites. With the integration of LCC, 58% of the 554 site-years show increased R² for GPP and ET estimations, 61% of the site-years show reduced biases for GPP estimation and 59% for ET estimation.
2. The temporal and spatial variability in LCC are equally important in improving flux estimations. The temporal variability in LCC is responsible for improved estimations in spring and autumn while the spatial variability in LCC explains more in summer. PFTs with strong seasonal cycles (i.e. DBF, CRO, GRA and SH) display major changes while evergreen species (i.e. ENF and EBF) display weak improvements in estimated GPP and ET after using spatially and temporally varying LCC in a model, which are likely attributed to the lack of leaf age representation in the LCC derivation algorithm.

3. LCC alone explains 30 (±13) % of the variability in GPP and 25 (±11) % of the variability in ET. The larger sensitivity of ET to abiotic factors explains the smaller improvements in ET than in GPP after using LCC.

This study demonstrates that satellite-derived LCC can be used to improve the modeled temporal and spatial variabilities in GPP and ET at the global scale. LCC can be considered as a necessary leaf physiological parameter within future TBMs to reduce the uncertainty in modelled fluxes and global carbon budgets.

4.7 Appendix

4.7.1 Validation of BEPS with and without LCC

Estimated GPP and ET from the BEPS using LCC-based Vcmax25 (BEPS-LCC) and the BEPS using fixed Vcmax25 (BEPS-REF) are validated with flux measurements from 124 sites in the FLUXNET2015 dataset. Correlation coefficient (R^2) and the root mean square error (RMSE) are used as the indicator to evaluate model performances.

4.7.1.1 Temporal validation

BEPS-LCC and BEPS-REF produce simulations in hourly steps. We aggregate the simulated fluxes to the daily scale and compared them with the flux data. Tables demonstrate the temporal validation of the estimated GPP and ET from BEPS-LCC and BEPS-REF.
Table 4-A1. Linear correlations between the daily estimated fluxes and measured fluxes. UNC refers to the observational uncertainties. Bold figures indicate higher $R^2$ or lower RMSEs. Figures in brackets are standard deviations.

<table>
<thead>
<tr>
<th>PFTs</th>
<th>GPP</th>
<th></th>
<th></th>
<th>ET</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BEPS-LCC</td>
<td>BEPS-REF</td>
<td>BEPS-LCC</td>
<td>BEPS-REF</td>
<td>BEPS-LCC</td>
<td>BEPS-REF</td>
</tr>
<tr>
<td></td>
<td>$N$</td>
<td>UNC (g/m$^2$/d)</td>
<td>$R^2$</td>
<td>RMSE (g/m$^2$/d)</td>
<td>$R^2$</td>
<td>RMSE (g/m$^2$/d)</td>
</tr>
<tr>
<td>CRO</td>
<td>12</td>
<td>0.23 (0.14)</td>
<td>0.61 (0.21)</td>
<td>3.72 (0.99)</td>
<td>0.54 (0.19)</td>
<td>4.33 (1.36)</td>
</tr>
<tr>
<td></td>
<td>DBF</td>
<td>0.35 (0.17)</td>
<td>0.85 (0.12)</td>
<td>2.41 (1.24)</td>
<td>0.78 (0.17)</td>
<td>2.68 (1.13)</td>
</tr>
<tr>
<td></td>
<td>EBF</td>
<td>0.80 (0.14)</td>
<td>0.41 (0.16)</td>
<td>3.37 (1.73)</td>
<td>0.37 (0.20)</td>
<td>3.28 (1.93)</td>
</tr>
<tr>
<td></td>
<td>ENF</td>
<td>0.26 (0.07)</td>
<td>0.71 (0.17)</td>
<td>2.29 (1.15)</td>
<td>0.74 (0.14)</td>
<td>2.23 (1.18)</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>0.40 (0.12)</td>
<td>0.79 (0.07)</td>
<td>2.17 (0.74)</td>
<td>0.81 (0.08)</td>
<td>2.16 (1.05)</td>
</tr>
<tr>
<td></td>
<td>GRA</td>
<td>0.24 (0.06)</td>
<td>0.61 (0.25)</td>
<td>1.85 (0.84)</td>
<td>0.60 (0.25)</td>
<td>2.13 (1.00)</td>
</tr>
<tr>
<td></td>
<td>SAV</td>
<td>0.27 (0.04)</td>
<td>0.53 (0.25)</td>
<td>2.60 (1.03)</td>
<td>0.44 (0.28)</td>
<td>2.88 (1.17)</td>
</tr>
<tr>
<td></td>
<td>SH</td>
<td>0.13 (0.05)</td>
<td>0.53 (0.31)</td>
<td>1.36 (0.72)</td>
<td>0.56 (0.35)</td>
<td>1.50 (1.12)</td>
</tr>
<tr>
<td></td>
<td>WET</td>
<td>0.15 (0.06)</td>
<td>0.70 (0.30)</td>
<td>2.02 (0.92)</td>
<td>0.62 (0.32)</td>
<td>1.97 (0.81)</td>
</tr>
<tr>
<td>Overall</td>
<td>124</td>
<td>0.38 (0.09)</td>
<td>0.65 (0.21)</td>
<td>2.38 (1.02)</td>
<td>0.63 (0.21)</td>
<td>2.54 (1.17)</td>
</tr>
</tbody>
</table>

Table 4-A1 shows that GPP estimates are improved by 6% in terms of $R^2$ and by 2% in terms of RMSE on average. After using LCC, 58% of the 554 site-years show increased $R^2$ for GPP validation while 58% of the site-years show increase $R^2$ for ET validation. With the exception of MF and ENF, other PFTs demonstrate smaller RMSE and larger $R^2$ when using BEPS-LCC. The PFTs that have the largest phenological cycles show the largest improvements in GPP estimates (e.g. DBF, WET and CRO), with $R^2$ value increasing by >0.07. RMSEs for DBF and CRO decrease by 0.61 g/m$^2$/day and 0.27 g/m$^2$/day, respectively. ET estimated by BEPS-LCC also shows some advantages over BEPS-REF: BEPS-LCC simultaneously increases $R^2$ and reduces the RMSEs for DBF.
and CRO; EBF, GRA, SAV, SH and WET demonstrates different degree of improvements with either higher $R^2$ or smaller RMSEs.

4.7.1.2 Spatial validation

We aggregated the fluxes simulated from BEPS-REF and BEPS-LCC to annual values. Estimated total annual GPP and ET of all sites are compared against the measured GPP and ET across the Fluxnet (Figure 4-A1).

![Graphs showing validation of annual GPP and ET](image)

**Figure 4-A1.** Validation of annual GPP (a, b) and ET (c, d) from BEPS-LCC (a, c) and BEPS-REF (b, d). There are 124 sites used for GPP validation and 110 sites used for ET.
Figure 4-A1 shows that BEPS-LCC is more capable of capturing the variability and magnitude of GPP than BEPS-REF. Estimated annual GPP from BEPS-LCC is correlated to observed GPP with $R^2=0.87$ and RMSE $=253.4$ g/m$^2$/yr, while for BEPS-REF GPP the validation results are $R^2= 0.81$ and RMSE $= 347.7$ g/m$^2$/yr. By using BEPS-LCC, the annual GPP estimation is improved by 6% in $R^2$ and the RMSE by 27%. Meanwhile, ET simulated by BEPS-LCC explains 34% of the measured ET variability while BEPS-REF explains 32%. The RMSE of estimated ET is also slightly reduced from 151.3 mm/yr in BEPS-REF to 148.3 mm/yr in BEPS-LCC. The average uncertainties of annual GPP and ET are 6.2 g/m$^2$/yr (1.4-24.3 g/m$^2$/yr) and 2.4 mm/yr (0.9-7.6 mm/yr), respectively, which are considerably smaller than the error of estimations (i.e. RMSEs). Therefore, the changes in $R^2$ and RMSEs made by LCC in Figure 4-A1 are credible and cannot be affected by observational uncertainties.

However, the changes made by LCC in estimated GPP and ET can vary with PFTs, because of the diversity in vegetation traits and local microclimate. Figure 4-A2 and Figure 4-A3 categorize site annual GPP and ET by PFT and plot them against the flux measurements in order to assess the influence of LCC in modelling fluxes for different ecosystems.
**Figure 4-A2.** Validation of site annual GPP for each PFT. Green dots are the results from BEPS-LCC, red dots are the results from BEPS. Top-equations denote the validation results of BEPS-LCC, bottom-equations denote those of BEPS-REF.

In Figure 4-A2, BEPS-LCC demonstrates advantages in GPP simulations for almost all PFTs with higher $R^2$ or smaller RMSEs. High $R^2$ and lower RMSE indicate the BEPS-LCC improved the simulation of spatial variations of fluxes, and satellite-derived LCC is a robust proxy of the spatial patterns of $V_{cmax25}$. Compare to the BEPS-REF results, the mean improvement in $R^2$ is 0.04, and the mean changes in RMSEs is -42 g/m$^2$/yr. Among nine PFTs, CRO shows the most significant improvement with $R^2$ increased by 0.19 and RMSE reduced by 232.1 g/m$^2$/yr. GRA, SAV and WET also show positive changes with higher $R^2$ and lower RMSE consistently. Forests types show complex and
less changes relative to the reference results. Only EBF, MF and SAV show higher annual GPP, while others show lower annual GPP.

**Figure 4-A3.** Validation of site annual ET for each PFT. Green dots are the results from BEPS-LCC, red dots are the results from BEPS. Top-equations denote the validation results of BEPS-LCC, bottom-equations describe those of BEPS-REF.

Figure 4-A3 shows that ET estimation of BEPS-LCC does not differ much from ET estimation of BEPS-REF in PFT-specific validation. Though ET from BEPS-REF and BEPS-LCC are both positively correlated with measurements, only CRO, GRA and SH show improvements in $R^2$ and RMSE simultaneously. The usage of site annual mean ET here erases the influence of LCC on the inter-annual variability in ET, therefore the improvements brought by LCC is not clear in this conventional validation. In addition,
the differences in the energy closure ratios of sites can affect the validation results, because part of the LE energies was not picked up by the eddy-covariance instruments [Foken, 2008] (Section 4.7.3).

4.7.2 Uncertainty of the modelling results

According to the structure of BEPS (Figure 4-2), the modelling uncertainties propagated from the inputs (i.e. metrological variables, LAI and CI) and model structure are almost the same for BEPS-REF and BEPS-LCC. When we investigated the systematic differences between BEPS-REF and BEPS-LCC, most modelling uncertainties can cancel each other, while the only source of uncertainty carried forward comes from satellite-derived LCC used by BEPS-LCC.

The uncertainty of satellite-derived LCC is incurred by errors of satellite-derived LAI and errors of MERIS surface reflectance. Early study on pre-processing of MERIS data suggested that the radiometric error is less than 2% between 400 nm and 900 nm, and less than 5% between 900 nm and 1050 nm [Curran and Steele, 2005]. This magnitude of error is not sufficient to detect variations of less 1 mg/m³ in chlorophyll content of water body. A recent study on ocean color also confirms uncertainty of MERIS reflectance is constrained at a scale of 5% in blue bands [Melin et al., 2016]. As for the LAI uncertainty, Verger et al. [2015] found the difference between MODIS or AVHRR LAI and GEOV1 LAI is less than 0.5 at more than 90% of the global pixels. Through a comparison with ground observations, the RMSEs between measured LAI and GEOV1 LAI is around 0.7 [Camacho et al., 2013]. In the coming uncertainty analysis, we designated four cases by changing the LAI by ± 0.7 and the reflectance by ±5% to derive LCC, running BEPS-LCC and repeating the validation process above. The purpose is to evaluate whether the changes in $R^2$ and RMSE found in spatial and temporal validation (Section 4.7.1.1) can be explained by modelling uncertainties.
Figure 4-A4. Uncertainties of the improvements brought by BEPS-LCC relative to BEPS-REF. (a-b) Spatial validation for GPP; (c-d) temporal validation for GPP; (e-f) spatial validation for ET; (g-h) temporal validation for ET.

In Figure 4-A4, $\Delta R^2$ larger than 0 and $\Delta$RMSEs smaller than 0 indicate that BEPS-LCC performs better than BEPS-REF. With the cases including uncertainties of satellite-derived LCC considered, BEPS-LCC shows higher $R^2$ and lower RMSE for GPP estimations than BEPS-REF in 78% and 73% of the cases, respectively. For ET, BEPS-LCC shows higher $R^2$ and lower RMSE than BEPS-REF in 62% and 60% of the cases, respectively. CRO, MF, SAV and SH are influenced most by uncertainties of LCC derivation. CRO is very sensitive to the LCC changes due to LAI, while SH is sensitive to the changes incurred by LAI and reflectance. $R^2$ and RMSEs of MF and SAV vary markedly among cases, largely because of the small dynamic range of their GPP and ET values and limited number sites for these two PFTs.

4.7.3 Energy closure problem
The energy closure problem of eddy-covariance measurements places some uncertainties in ET validation. In theory, measured net radiation, soil heat flux (G), sensible heat (H) and LE at a site should converge to the energy balance: \( Rn - G = LE + H \). However, flux towers find that the sum of LE and H picked up by instruments is smaller than the residue of \( Rn-G \). This issue is commonly attributed to the uncertainties in the energy storage of vegetation [e.g. Lindroth et al., 2010] and the big eddies missed by the eddy-covariance instruments [e.g. Foken, 2008]. A protocol for correcting the energy closure ratio is still underway for eddy-covariance measurements [Gao et al., 2017]. Figure 4-A5 compiles the energy closure ratio for each PFT using the FLUXNET2015 dataset.

According to Figure 4-A5, the average energy closure ratios distribute between 76% and 85% for each PFT, while the inter-sites variations range between 9% and 24%. Past reviews found a very similar energy closure ratio around 80% across FLUXNET [Wilson et al., 2002; Wang and Dickinson, 2012]. Given the available energy closure ratio, it is reasonable to foresee that the bias of modelled ET tends to be positive since LE measurements miss part of the energy. To avoid this known bias in ET validation, FLUXNET2015 provides a dataset column named LE_CORR as an alternative. In its data pipeline, it assumes that the Bowen ratio keeps fixed during a period of \( \pm 5 \) hours, and the missing energy is divided between LE and H according to the Bowen ratio. However, using LE_CORR to validate BEPS-LCC modelled ET does not show better \( R^2 \)
and RMSE than the results shown previously. The biases of ET estimations during the growing seasons are reduced but the biases in non-growing seasons increased. It implies that using the Bowen ratio to correct LE is a viable solution to adjust total ET, some seasonal variabilities should be added to the constant Bowen ratios to improve the partition of missing energies.

4.8 References
Camacho, F., J. Cernicharo, R. Lacaze, F. Baret, and M. Weiss (2013), GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products,


Ehler, G. J., N. J. Livingston, D. L. Harrison, T. A. Black, and J. A. Moran (2006), Low


Jiang, C., and Y. Ryu (2016), Multi-scale evaluation of global gross primary productivity


Liu, R., R. Shang, Y. Liu, and X. Lu (2017), Global evaluation of gap-filling approaches for seasonal NDVI with considering vegetation growth trajectory, protection of key


Ryu, Y. et al. (2011), Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales, Global Biogeochemical Cycles, 25(4), n/a-n/a, doi:10.1029/2011GB004053.


Simic, A., J. M. Chen, and T. L. Noland (2011), Retrieval of forest chlorophyll content


Wu, J. et al. (2016), Leaf development and demography explain photosynthetic seasonality
in Amazon evergreen forests, *Science*, 351(6276), 972 LP-976.


Zhang, Y. et al. (2016), Precipitation and carbon-water coupling jointly control the interannual variability of global land gross primary production, *Scientific Reports*, 6(August), 39748, doi:10.1038/srep39748.

Chapter 5
Summary

5.1. Lessons learned from this research

An accurate representation of leaf structural and physiological status is critical to the simulation of transpiration – the major component of terrestrial evapotranspiration (ET). Leaf structural traits such as leaf area index (LAI) and clumping index (CI) describe the distribution and area of leaves in a canopy, and thus can be used to simulate the light environment inside the canopy and then work as a basis to upscale the estimated carbon-water fluxes from the leaf level to the ecosystem scale. Leaf physiological traits such as leaf chlorophyll content (LCC) indicate the amount of pigments and enzymes supporting the biochemical processes in leaves; thus, the variability of LCC can add seasonality to photosynthesis rate ($A_n$) and stomatal conductance ($g_s$) which regulate the water-carbon exchange between vegetation and the atmosphere. Since water and carbon fluxes are tightly coupled through stomata – the small pores on leaf surfaces, this thesis also has significant impact on the quantification of global carbon uptake (gross primary productivity; GPP). The results suggest that:

1. A two-leaf scheme (TL) that exploits LAI and CI to separate a canopy into sunlit and shaded leaves would considerably improve ET and GPP estimations relative to the commonly used big-leaf scheme (BL) over nine flux towers across Canada. BL underestimates ET and GPP at all the sites and its underestimation of fluxes increases with increasing LAI. Another popular upscaling scheme – the two-big-leaf scheme (TBL) – was also compared. Though TL and TBL both consider the separation of sunlit and shaded leaf in canopies, TBL underestimates ET by 7% and shows increasing biases in estimated ET and GPP with increasing LAI as well, due to the uncertainty incurred from the artificial calculation of canopy conductance ($G_c$) and other biochemical parameters in TBL. The inadequate performance of BL and TBL shown in this chapter invalidates the big leaf assumption underlying the Penman-Monteith equation for ET.
calculations in process-based models that follow the principle of carbon-water coupling. Therefore, the TL scheme is regarded as the optimal upscaling scheme for estimating ET. Although the difference in simulated ET between TL and TBL models is small in most cases, it is important to recognize that the TL model structure is compatible with the physics of water flow in the canopy and avoids a significant error in the TBL model structure without incurring additional computation.

2. LCC is a direct proxy for the leaf photosynthetic capacity (maximum carboxylation rate; Vcmax25) of trees based on measurements from the Borden forest site. By incorporating LCC-based Vcmax25 in a terrestrial biosphere model – Boreal Ecosystem Productivity Simulator (BEPS), the estimates of GPP and ET are both improved at the site, especially during the transitional periods in spring and autumn, compared to three commonly used Vcmax25 parameterization methods. In addition, by incorporating LCC in the two-leaf scheme of BEPS, this study is able to quantify the physiological difference between sunlit and shaded leaves. The variations in light environments drive sunlit leaves to have a higher Vcmax25 to LCC ratio than shaded leaves to maximize the total carbon uptake of the canopy.

3. Satellite-derived LCC represents the spatial and temporal variability in Vcmax25 over 124 eddy-covariance sites (554 site-years) across nine plant functional types over the globe. By using satellite-derived LCC in BEPS, most of site-years show reductions in bias of GPP and ET estimates. The plant functional types with strong seasonal cycles (i.e. cropland, deciduous broadleaf forest, grasslands) show bigger improvements than evergreen species. LCC mainly improves fluxes estimates through considering the spatial variability of Vcmax25. In spring and autumn, the enhanced seasonality in Vcmax25 caused by LCC also drives the improvements in GPP and ET estimates.

5.2. Topics for further research

LAI is an established structural parameter used in most TBMs to quantify the dynamics of vegetation. CI is then added in some TBMs to achieve the simulation of instantaneous radiation on leaves through separating sunlit and shaded leaves. LAI and CI form the
basis of the upscaling schemes in TBM to upscale leaf-level fluxes to the canopy scale. This thesis aims at assessing three mainstream upscaling schemes and answering the unsolved questions about which one is the best for simulating ET. Considering the efficacy and efficiency of models, the two-leaf scheme is shown to be the optimal choice over nine eddy-covariance towers in Canada. In contrast to current computational capacities, the two-leaf scheme simply regards the mean solar radiation over a leaf group as the solar radiation on individual leaves in that group. This simplification might have neglected some variations of radiation inside canopies, especially for shaded leaves. Further studies can focus on the development of a more realistic multi-layer sunlit and shaded scheme, pending on the advances in computational capacities. This thesis also designs a framework to incorporate LCC into the two-leaf scheme which could be upgraded if we had more ground measurements in the future.

LCC is shown to be a robust proxy for Vcmax25 in TBM, and this thesis demonstrates the possibility of using satellite-derived LCC for large-scale simulation of GPP and ET. This thesis encourages more research on utilizing LCC or other leaf physiological parameters in models. As Chapter 3 and Chapter 4 have demonstrated, the uncertainties in the Vcmax25-LCC relationships should be the focus for further improvements, since it has been recognized that Vcmax25-LCC relationships vary with leaf light environments, species and possibly leaf age while the measurements to quantify these variations are insufficient.

The trend of ET in the context of climate change is subject a wide range of physical and biotic factors. For example, the rising temperature and enhanced precipitation can intensify ET, while droughts and a dimming earth can reduce the ET rate. The influence of the physical factors is intertwined with biotic factors by stimulating the changes in biodiversity and the adaptions of plants. A warmer climate drives the poleward expansion of broadleaf trees that usually have higher transpiration rates than conifers. Some species can also adjust their water loss with their photosynthetic rate, which is incurred by the increasing atmospheric CO₂ concentration and changes in the nitrogen availability. As an indicator for the photosynthetic nitrogen pool, the LCC can provide a framework to explore the effects of biotic factors on the future ET.
We also cannot help but find that GPP is improved more than ET in studies in Chapters 2, 3 and 4. We analysed the sensitivities of measured GPP and ET to a diverse range of biotic and abiotic factors in Chapter 4. The results show that abiotic factors are more influential on ET while GPP is more determined by biotic factors. ET improvements are most obvious in closed or high-LAI canopies where transpiration contributes to the majority of ET. However, in open canopies where transpiration is not the dominant water flux, evaporation from the soil surface can make use of the radiation missed by leaves and markedly influence the total water flux. Therefore, further studies on constraining the uncertainty of ET need to focus on the simulation of evaporation from the soil surface and wet leaves. My thesis research uses a simple model to estimate evaporation components, and the model may also need improvement based on other data sources. Moreover, evaporation from the ground surface is a strong determinant of soil moisture which is a known regulator for stomatal conductance. Thereby, an accurate evaporation will also help in accurately simulating transpiration.