Reducing Uncertainty in The Biosphere-Atmosphere Exchange of Trace Gases

by

Kimberly Novick

Department of Environment
Duke University

Date: ____________________________

Approved:

______________________________
Gabriel Katul, Advisor

______________________________
Ram Oren

______________________________
Amilcare Porporato

______________________________
Hans Peter Schmid

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Environment in the Graduate School of Duke University 2010
ABSTRACT
(Environmental Science)

Reducing Uncertainty in The Biosphere-Atmosphere Exchange of Trace Gases

by

Kimberly Novick

Department of Environment
Duke University

Date: _______________________

Approved:

__________________________
Gabriel Katul, Advisor

__________________________
Ram Oren

__________________________
Amilcare Porporato

__________________________
Hans Peter Schmid

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Environment in the Graduate School of Duke University 2010
Abstract

Much of the anthropogenic emissions of greenhouse gases (GHGs) is cycled through the terrestrial biosphere, thereby moving the topic of quantifying the exchange rates between the terrestrial biosphere and the atmosphere of these gases from the margins of traditionally disjointed disciplines such as hydrology, micrometeorology, and plant physiology to a unifying and central research topic in climate science. These fluxes are governed by biophysical processes such as photosynthesis, transpiration, and microbial respiratory processes that have multiple and dynamic drivers such as meteorology, disturbance regimes, and long-term land cover and climate change. These complex processes occur over a broad range of temporal (seconds to decades) and spatial (millimeters to kilometers) scales, complicating efforts to quantify biosphere-atmosphere exchanges over the diverse range of global biomes and necessitating the application of simplifying models to forecast fluxes at the regional and global scales required by climate mitigation and adaptation policymakers.

Over the long history of biophysical research, much progress has been made towards developing appropriate models for the biosphere-atmosphere exchange of GHGs. Many processes are well represented, particularly at the leaf scale. However, many remain poorly understood despite several decades of field measurements, and models still do not perform robustly over coarse spatial scales and long time frames. Indeed, model and parameter uncertainty remains a major contributor to difficulties in constraining the atmospheric budgets of greenhouse gases.
The central objective of this dissertation is to reduce uncertainty in the quantification and up-scaling of the biosphere-atmosphere exchange of greenhouse gases. Naturally, addressing all aspects of this problem remains well beyond the scope of a single dissertation. The compass of this work is to strategically address key research questions through five case studies. In Chapter 1, nocturnal evapotranspiration – a physiological process that had been largely ignored until recent years which is quantified and modeled in three unique ecosystems co-located in central North Carolina, U.S.A. In the second chapter, long-term drivers of evapotranspiration are explored by developing and testing theoretical relationships between plant water use and hydraulic architecture that may be readily incorporated into terrestrial ecosystem models. The third chapter builds on this work by linking key parameters of carbon assimilation models to structural and climatic indices that are well-specified over much of the land surface in an effort to improve model parameterization schemes. The fourth chapter directly addresses questions about the interaction between physiological carbon cycling and disturbance regimes in current and future climates, which are generally poorly represented in terrestrial ecosystem models. And the last chapter explores effluxes of methane and nitrous oxide (which are historically understudied) in addition to CO₂ exchange in a large temperate wetland ecosystem (that is a historically understudied biome). While these five case studies are somewhat distinct investigations in terms of processes and the nature of up-scaling, they are all grounded in the principles of biophysics, are all generally conducted to address exchange rates at the ecosystem scale, and all rely on similar measurement and mathematical modeling techniques.
# Contents

Abstract iv

List of Tables xii

List of Figures xiv

List of Abbreviations and Symbols xvii

Acknowledgements xxiv

1 Introduction 1

1.1 Trace gas exchange between the biosphere and the atmosphere - governing processes and complexities 4

1.2 The need for models, and a summary of successes and challenges 8

1.2.1 Biosphere-atmosphere interactions - some research and modeling successes 11

1.2.2 Biosphere-atmosphere interactions - some remaining challenges 12

1.3 Some relatively new research tools 14

1.3.1 The eddy covariance technique 15

1.3.2 Remotely-sensed vegetation indices 16

1.3.3 Advanced analytical methods 17

1.4 The case studies 18

2 Nocturnal evapotranspiration in eddy covariance records from three co-located ecosystems in the Southeastern U.S.: Implications for annual fluxes 23

2.1 Introduction 23
2.6.4 Influence of $T$ and $E$ on $ET_N$ ........................................ 52
2.6.5 Evaluating gapfilling methods for $ET_N$ ................................. 53
2.6.6 Broader implications and conclusions ................................. 55

3 The relationship between reference canopy conductance and simplified hydraulic architecture\(^1\) ........................................ 57

3.1 Introduction ........................................................................ 57
3.2 Theoretical considerations and hypotheses ......................... 58
  3.2.1 Relating transpiration and conductance to hydraulic architecture 58
  3.2.2 Separating fast and slow responses .............................. 59
3.3 Methods ............................................................................ 64
  3.3.1 Survey 1 – relationships between $G_{Sref}$, $h$, and $A_S/A_L$ .......... 64
  3.3.2 Allometric equations for $A_S/A_L$ .................................. 69
  3.3.3 Statistical tests and optimization .................................. 69
3.4 Results .............................................................................. 69
  3.4.1 Changes in $G_{Sref}$ with $A_S/A_L$ and $h$ ............................ 69
  3.4.2 Relationship between $A_S/A_L$ and $h$ ............................ 75
3.5 Discussion ........................................................................ 77
  3.5.1 The hydraulic controls on stomatal conductance across species 77
  3.5.2 Mechanisms and limits to hydraulic compensation within species 80
  3.5.3 Variation in the rate of change of $A_S/A_L$ ....................... 82
  3.5.4 Broader implications for ecosystem-to-regional scale carbon and water cycle modeling ............................ 84

4 A characterization of the variability in ecosystem-scale model parameters using Bayesian inversion of eddy covariance data ............................. 86

4.1 Introduction ........................................................................ 86
4.2 Model selection .................................................................... 90
  4.2.1 The non-rectangular hyperbola model for NEE ............... 91
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.2</td>
<td>A Jarvis-type model for canopy stomatal conductance</td>
<td>91</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Linking CO₂ and H₂O exchange</td>
<td>93</td>
</tr>
<tr>
<td>4.3</td>
<td>Methods</td>
<td>94</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Eddy covariance and meteorological data – sources and quality control</td>
<td>94</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Ancillary data</td>
<td>95</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Additional data filtering and study period definition</td>
<td>97</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Error in the eddy covariance measurements</td>
<td>98</td>
</tr>
<tr>
<td>4.3.5</td>
<td>Bayesian inversion</td>
<td>98</td>
</tr>
<tr>
<td>4.3.6</td>
<td>Determining water use efficiency</td>
<td>101</td>
</tr>
<tr>
<td>4.4</td>
<td>Results</td>
<td>101</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Data availability</td>
<td>101</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Random measurement error</td>
<td>102</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Bayesian inversion - posterior PDFs</td>
<td>103</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Parameter variation across PFTs, within PFTs, and with time</td>
<td>106</td>
</tr>
<tr>
<td>4.4.5</td>
<td>Parameter variation across PFTs and with canopy architecture</td>
<td>108</td>
</tr>
<tr>
<td>4.4.6</td>
<td>The relationship between λₜ and D</td>
<td>110</td>
</tr>
<tr>
<td>4.5</td>
<td>Discussion</td>
<td>112</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Assimilation of eddy covariance data with Bayesian inversion techniques</td>
<td>117</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Sources of error: Random measurement error, model selection, and LAI</td>
<td>118</td>
</tr>
<tr>
<td>4.5.3</td>
<td>Inter-specific relationships between Gₜref, β, γ and structural variables</td>
<td>120</td>
</tr>
<tr>
<td>4.5.4</td>
<td>The sensitivity of stomatal conductance and water use efficiency to D</td>
<td>122</td>
</tr>
<tr>
<td>4.5.5</td>
<td>Issues in up-scaling from leaf to canopy</td>
<td>123</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Summary of measured greenhouse gas exchange rates for wetland ecosystems</td>
<td>9</td>
</tr>
<tr>
<td>1.2</td>
<td>Key features of select Dynamic Global Vegetation Models</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>Comparison of $ET_N$ gapfilling methods</td>
<td>35</td>
</tr>
<tr>
<td>2.2</td>
<td>Evapotranspiration flux data remaining after quality control filtering</td>
<td>37</td>
</tr>
<tr>
<td>2.3</td>
<td>Annual and growing season gapfilled evapotranspiration estimates</td>
<td>44</td>
</tr>
<tr>
<td>2.4</td>
<td>Statistical estimates of performance for the gapfilling methodologies</td>
<td>45</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of studies used to assess the relationship between $G_{Sref}$, height, and $A_S/A_L$</td>
<td>67</td>
</tr>
<tr>
<td>3.2</td>
<td>Summary of studies used to assess the relationship between sapwood-to-leaf area ratio and mean canopy height</td>
<td>70</td>
</tr>
<tr>
<td>3.3</td>
<td>The relative change in $G_{Sref}$, $h$, and $A_S/A_L$ for select ecosystems</td>
<td>81</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of Ameriflux sites included in the Bayesian inversions</td>
<td>96</td>
</tr>
<tr>
<td>4.2</td>
<td>The range of the uniform priors</td>
<td>100</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison of model-data and random measurement errors</td>
<td>103</td>
</tr>
<tr>
<td>4.4</td>
<td>Means and 95% confidence intervals for the posterior distributions</td>
<td>105</td>
</tr>
<tr>
<td>4.5</td>
<td>Coefficients of variation of model parameters for cross-$PFT$, within-$PFT$, and inter-annual comparisons</td>
<td>110</td>
</tr>
<tr>
<td>5.1</td>
<td>ANOVA table for the impacts of elevated CO$_2$, fertilization, size class, and sampling date on resin flow</td>
<td>139</td>
</tr>
<tr>
<td>6.1</td>
<td>Quality control thresholds for the scalar-gradient greenhouse gas flux estimates</td>
<td>156</td>
</tr>
</tbody>
</table>
6.2 Available data for greenhouse gas flux estimates from the restored wetland .......................................................... 160

6.3 Measurement error associated with the gas chromatographer ......................................................... 176
# List of Figures

1.1 The drivers of the biosphere-atmosphere exchange of CO$_2$ and H$_2$O, and the temporal and spatial scales on which they act .......................... 5
1.2 An eddy covariance tower in the Duke Forest ................................. 16
1.3 Ratios of nocturnal to daytime transpiration in recent studies ........... 19
2.1 Frequency of observed $ET_N$ .................................................. 38
2.2 Seasonal patterns in observed $ET_N$ ........................................... 39
2.3 The relationship between $ET_N$ and meteorological drivers .......... 42
2.4 An evaluation of the process-based model for nocturnal $ET$ ............ 43
2.5 Relative frequency distributions of observed and gapfilled $ET_N$ ...... 46
2.6 Nocturnal leaf boundary layer and aerodynamic conductance ......... 52
3.1 Errors related to theoretical assumptions linking $G_{Sref}$ to $h$ .......... 61
3.2 The inter-specific relationship between $G_{Sref}$, $A_S/A_L$, and $h$ ...... 72
3.3 Relationships between $G_{Sref}$ and $1/h$ within biomes .................. 73
3.4 Relationships between $G_{Sref}$ and LAI ....................................... 75
3.5 Variation in $G_{Sref}$ explained by variation in $\alpha_S$ .................. 77
3.6 Variation in $\alpha_S$ as a function of mean annual temperature ....... 78
4.1 The central moments of the posterior distributions for $NEE$ and $G_T$ model parameters ............................................................. 107
4.2 Correlation between posterior distributions ................................. 108
4.3 The range of parameter values with $PFTs$ .................................... 109
4.4 Inter-annual variability in posterior distributions ................. 111
4.5 Relationships between model parameters and canopy architecture . 112
4.6 Relationships between model parameters and LAI ..................... 114
4.7 Relationships between water use efficiencies and $D$ for short and long time scales ........................................... 115
4.8 Canopy assimilation as a function of light for a single layer and multi-layer canopy, and as a function of temperature ..................... 124
4.9 A comparison of a PFT-based parameterization schemes to a height-based parameterization scheme ................................. 127
5.1 Time series of resin flow, $ANPP_W$, and gross canopy assimilation ........ 136
5.2 Relationship between resin flow and $ANPP_W$ .......................... 137
5.3 Relationship between resin flow and native nitrogen availability ....... 140
6.1 The Timberlake Restoration Project ........................................ 152
6.2 Meteorological conditions at the TRP during the study period ....... 158
6.3 Wind direction and the flux footprint .................................. 161
6.4 The relationship between eddy-covariance derived CO$_2$ fluxes and light 163
6.5 Energy fluxes in the Timberlake Restoration Project .................. 164
6.6 GHG fluxes from the scalar gradient methodology ...................... 166
6.7 The impact of an artificial drawdown experiment on CO$_2$ and CH$_4$ fluxes ................................................... 167
6.8 Scalar gradient flux estimates derived from unique subsample sets ... 169
6.9 The relationship between surface conductance ($G_T$) and vapor pressure deficit ($D$) for a dryland and an inundated area in the Timberlake Restoration Project .......................................... 174
6.10 A comparison of measured and modeled eddy diffusivity .............. 177
B.1 Posterior distribution for Size, CO$_2$, and Fertilization effects ........ 195
B.2 Posterior distribution for random individual effects and random error 196
B.3 The month effects ....................................................... 197
C.1 GS flux error resulting from concentration measurement error, as a function of the number of measurements .............................. 199
# List of Abbreviations and Symbols

**Latin Symbols**

- $A$  
  Gross CO$_2$ assimilation rate
- $a$  
  An empirical parameter of a process-based model for nocturnal evapotranspiration that scales with conductance
- $A_L$  
  Leaf area
- $A_S$  
  Sapwood area
- $a_V$  
  A temperature dependent constant to infer transpiration from stomatal conductance and vapor pressure deficit
- $A_z$  
  Slope parameter for log-log relationship between a scalar quantity and vertical displacement
- $a_{CO_2}$  
  Ambient CO$_2$
- $AF$  
  Fertilized study plot in ambient CO$_2$
- $ANPP_W$  
  Aboveground net wood primary production
- $AU$  
  Unfertilized study plot in ambient CO$_2$
- $b$  
  An empirical parameter of a process-based model for nocturnal evapotranspiration that scales with the sensitivity of conductance to vapor pressure deficit
- $b'$  
  Intercept parameter in the Ball-Berry conductance model
- $b_d$  
  An empirical parameter describing sensitivity of stomatal conductance to vapor pressure deficit
- $B_z$  
  Intercept parameter for log-log relationship between a scalar quantity and vertical displacement
- $BE$  
  Bias error
\(c\) A temperature dependent constant to infer evaporation from soil resistances

\(c_1\) One of two empirical constants describing the sensitivity of stomatal conductance to light

\(c_2\) One of two empirical constants describing the sensitivity of stomatal conductance to light

\(c_a\) Atmospheric CO\(_2\) concentration

\(C_d\) Drag coefficient

\(c_d\) Slope parameter for the relationship between \(c_i/c_a\) and vapor pressure deficit

\(c_i\) Intercellular CO\(_2\) concentration

\(c_p\) Specific heat capacity for dry air

\(C.V.\) Coefficient of variation

\(CDIAC\) Carbon Dioxide Information Analysis Center

\(D\) Vapor pressure deficit

\(d\) Zero plane displacement

\(d_d\) Intercept parameter for the relationship between \(c_i/c_a\) and vapor pressure deficit

\(d_l\) Characteristic leaf dimension

\(DAYNIGHT\) A gapfilling model that relates nocturnal to daytime evapotranspiration

\(DBH\) Diameter at breast height

\(DGVM\) Dynamic Global Vegetation Model

\(DISTSAMP\) Marginal distribution sampling

\(DS\) A dataset

\(E\) Evaporation

\(EC\) Eddy covariance

\(eCO_2\) elevated CO\(_2\)
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF</td>
<td>Fertilized study plot in elevated CO₂</td>
</tr>
<tr>
<td>ET</td>
<td>Evapotranspiration</td>
</tr>
<tr>
<td>ETₐ</td>
<td>Daytime evapotranspiration</td>
</tr>
<tr>
<td>ETₙ</td>
<td>Nocturnal evapotranspiration</td>
</tr>
<tr>
<td>EU</td>
<td>Unfertilized study plot in elevated CO₂</td>
</tr>
<tr>
<td>F&lt;sub&gt;CH₄&lt;/sub&gt;</td>
<td>Net ecosystem flux of CH₄</td>
</tr>
<tr>
<td>F&lt;sub&gt;CH₄,SG&lt;/sub&gt;</td>
<td>Scalar-gradient derived net ecosystem flux of CH₄</td>
</tr>
<tr>
<td>F&lt;sub&gt;CO₂&lt;/sub&gt;</td>
<td>Net ecosystem flux of CO₂, equivalent to NEE</td>
</tr>
<tr>
<td>F&lt;sub&gt;CO₂,EC&lt;/sub&gt;</td>
<td>Eddy-covariance derived net ecosystem flux of CO₂</td>
</tr>
<tr>
<td>F&lt;sub&gt;CO₂,SG&lt;/sub&gt;</td>
<td>Scalar-gradient derived net ecosystem flux of CO₂</td>
</tr>
<tr>
<td>F&lt;sub&gt;N₂O&lt;/sub&gt;</td>
<td>Net ecosystem flux of N₂O</td>
</tr>
<tr>
<td>F&lt;sub&gt;N₂O,SG&lt;/sub&gt;</td>
<td>Scalar-gradient derived net ecosystem flux of N₂O</td>
</tr>
<tr>
<td>FACE</td>
<td>Free-Air Carbon Enrichment</td>
</tr>
<tr>
<td>G</td>
<td>Tree growth</td>
</tr>
<tr>
<td>G&lt;sub&gt;aero&lt;/sub&gt;</td>
<td>Canopy aerodynamic conductance</td>
</tr>
<tr>
<td>G&lt;sub&gt;BL&lt;/sub&gt;</td>
<td>Leaf boundary layer conductance</td>
</tr>
<tr>
<td>G&lt;sub&gt;H&lt;/sub&gt;</td>
<td>Ground heat flux</td>
</tr>
<tr>
<td>G&lt;sub&gt;S&lt;/sub&gt;</td>
<td>Stomatal conductance to water vapor per unit leaf area</td>
</tr>
<tr>
<td>G&lt;sub&gt;Sref&lt;/sub&gt;</td>
<td>Stomatal conductance to water vapor per unit leaf area at a reference environmental state</td>
</tr>
<tr>
<td>G&lt;sub&gt;T&lt;/sub&gt;</td>
<td>Stomatal conductance to water vapor per unit ground area</td>
</tr>
<tr>
<td>G&lt;sub&gt;Tref&lt;/sub&gt;</td>
<td>Stomatal conductance to water vapor per unit ground area at a reference environmental state</td>
</tr>
<tr>
<td>g</td>
<td>gravitational acceleration</td>
</tr>
<tr>
<td>GC</td>
<td>Gas chromatographer</td>
</tr>
<tr>
<td>GCM</td>
<td>General Circulation Model</td>
</tr>
<tr>
<td>GHG</td>
<td>Greenhouse gas</td>
</tr>
</tbody>
</table>
Growing season
Sensible heat flux
Canopy height
Intergovernmental Panel on Climate Change
Leaf-level hydraulic conductivity
Eddy diffusivity for heat
Inferred eddy diffusivity from measured sensible heat flux and temperature profiles
Modeled eddy diffusivity for heat in neutral conditions
Modeled eddy diffusivity for heat that includes stability corrections
Eddy diffusivity of scalar $s$
Tissue-specific hydraulic conductivity
A likelihood function
Leaf area index
Latent Heat Flux
Light Detection and Ranging
The sensitivity of stomatal conductance to vapor pressure deficit
Slope parameter in the Ball-Berry conductance model
resin flow
reference resin flow
Mean absolute error
Metropolis-Hastings
Mobile Micro-meteorological Station
Multiple imputation
Net ecosystem exchange of CO$_2$
A probability function
$P_M$ Mean annual precipitation

$PAR$ Photosynthetically active radiation

$PDF$ Probability density function

$PDSI$ Palmer Drought Severity Index

$PFT$ Plant functional type

PROCESS A gapfilling model based on multiple non-linear regression

$R_n$ Net radiation

$r_{s,a}$ Sum of surface and aerodynamic soil resistances

$RH$ Relative humidity

$RRMSE$ Relative root mean square error

$s$ Instantaneous scalar concentration

$SG$ The scalar gradient methodology for inferring scalar fluxes

$T$ Transpiration per unit ground area

$T_a$ Air temperature

$T_m$ Mean annual air temperature

$t_p$ Time since a rain event

$T_r$ Transpiration per unit leaf area

$TRP$ Timberlake Restoration Project

$TWR_A$, $TWR_B$ Tower A and Tower B

$u$ Mean wind speed vector

$u$ Wind speed in the x-direction

$\bar{u}$ Mean horizontal wind speed

$u^*$ Friction velocity

$u_{sub}^*$ Friction velocity at the canopy floor

$v$ Wind speed in the y-direction

$v_c$ Molecular diffusivity of a scalar
\( w \) Wind speed in the z-direction

\( WD \) Wind direction

\( \mathbf{X}_{it} \) Vector of fixed effects

\( x_p \) Peak of the source-weight function for the flux footprint

\( z_{MEAS} \) Measurement height

\textsc{zero} A simple filter that sets all missing data to 0

**Greek symbols**

\( \alpha \) mean apparent quantum yield

\( \alpha_R \) Vector of regression parameters

\( \alpha_{R,S} \) Regression parameter for size class

\( \alpha_{R,C} \) Regression parameter for CO\(_2\) treatment

\( \alpha_{R,C \times N} \) Regression parameter for the CO\(_2\) and N interaction

\( \alpha_{R,N} \) Regression parameter for N treatment

\( \alpha_S \) slope of the relationship between sapwood-to-leaf area and height

\( \beta \) Light-saturated CO\(_2\) assimilation rate

\( \beta_{R,i} \) Random individual effects

\( \beta_S \) Intercept of the relationship between sapwood-to-leaf area and height

\( \gamma \) Ecosystem respiration

\( \gamma_1 \) Skewness

\( \gamma_2 \) Kurtosis

\( \delta \) Random measurement error in eddy covariance records

\( \varepsilon \) Random error in a Gaussian model, equivalent to the standard deviation

\( \varepsilon_{GT} \) Random error in total canopy conductance to water vapor

\( \varepsilon_{NEE} \) Random error in the net ecosystem exchange of CO\(_2\)
\( \varepsilon_{R.it} \) Random error in the Bayesian resin flow model

\( \zeta \) Atmospheric stability parameter

\( \Theta \) A set of parameters

\( \theta \) Soil moisture content

\( \theta_f \) Curvature of the relationship between net ecosystem exchange of CO\(_2\) and light

\( \kappa \) von Karman constant

\( \kappa_{R.t} \) Fixed month effects

\( \lambda \) Marginal water use efficiency

\( \mu \) The mean of a sample

\( \rho_w \) density of water

\( \sigma(\cdot) \) standard deviation

\( \sigma^2_R \) Variance parameter for random effects

\( \tau^2_R \) Variance parameter for individual effects

\( \phi(\zeta) \) Atmospheric stability correction

\( \psi_{leaf} \) Leaf water potential

\( \psi_{soil} \) Soil water potential

Averaging notation

\( \langle \cdot \rangle \) Denotes ensemble averaging

\( \langle \cdot \rangle \) Denotes Reynolds averaging over turbulent time scales

\( \cdot \) Primed quantities denote variations about the mean state
Acknowledgements

I would like to thank the National Science Foundation (through their Graduate Research Fellowship Program) and the James B. Duke Fellowship Program, which both provided stipend and tuition support for my graduate studies. Support for field work in the Duke Forest was provided by the U.S. Department of Energy (DOE) through the Office of Biological and Environmental Research (BER) Terrestrial Carbon Processes (TCP) program (Grants #10509-0152, DE-FG02-00ER53015, and DE-FG02-95ER62083), and the United States-Israel Binational Agricultural Research and Development (BARD, Research Grant No. IS3861-06). Field work in the Timberlake Restoration Project was supported by the US Department of Energy’s National Institute for Climate Change Research.

I am very grateful for the instruction I have received from my advisor Gaby Katul, who is undoubtedly one of the most generous mentors in the field. I am also indebted to the other members of my committee, including Ram Oren (who, at times, seems to have spent as much time advising me as his own students), Amilcare Porporato, and Hans Peter Schmid, for their insightful comments and suggestions. I would also like to acknowledge Emily Bernhardt for her advisement on the Timberlake field project.

I would like to thank a number of individuals for providing indispensable field and technical support. These include members of the Katul Lab (Mario Siqueira, Paul Stoy, Jehn-Yih Juang, Sally Thompson, and Alexandra Konings), the Oren Lab
(Eric Ward, Heather McCarthy, Christopher Oishi, Jeff Pippen, Hyun Kim, and Yavor Parashkevov), the Bernhardt Lab (Jennifer Morse, Medora Burke-Scoll, Marcelo Ardon, Jill-Thompson Greiner, Brooke Hassett, and Rebecca Sanders-DeMott), the Hain lab at N.C. State (who provided instruction and equipment for the resin experiment). Thanks also to the FACE site and Duke Forest staff, including Judd Edeburn, Robert Nettles, and David Cooley. I would especially like to thank Andrea Casanova and Javier Lascurain, Jr. for assisting with data collection during their tenure as undergraduate and masters students, respectively. I also would like acknowledge Bud Needham, the Great Dismal Swamp Mitigation Bank, and the fine people of Columbia, N.C. (including the nice farmer who used what has to be the world’s largest tractor to extract our Mobile Meteorological Station from a giant pit of mud).

The Nicholas School staff, including Nancy Morgans and Meg Stephens, should also be noted for their excellent organization support.

And finally, I would like to acknowledge my family. A million thanks to my parents, Chris and Andy Novick, who have unwaveringly supported every aspect of my education, from diligently checking my first-grade math homework to paying my college tuition to providing babysitting services during the final days of my tenure as a graduate student. Thanks to my brother and sister, Paul and Emily, for the lively dinner-table math and science debates. Thanks to my partner, Mike Jackson, who patiently endured many attempts on my part to explain my latest research challenges and frustrations using Sharpies and bar napkins. And last but not least, thanks to John Shepard Jackson for his excellent napping skills.
Introduction

In the 2007 Assessment Report on Climate Change, the Intergovernmental Panel on Climate Change (IPCC) asserts that the evidence for a warming climate is now "unequivocal" (IPCC, 2007a). Further, the panel states with high confidence that the strongest driver of this warming trend is the anthropogenic emissions of greenhouse gases (GHGs, including CO$_2$, CH$_4$, N$_2$O, and halocarbons), which are heat trapping gases that determine the radiative forcing of the atmosphere. Given the sustained emissions rates of these gases together with their long atmospheric residence times, this warming trend is expected to continue throughout the 21st century with a global temperature increase of 1.8 to 4.0 °C by 2099 (Meehl et al., 2007).

The expected impacts of a warmer world on human societies and natural ecosystems are diverse and complex, and have been well reviewed elsewhere (Parry et al., 2007; Houghton, 2004). Briefly, some of the most widely discussed impacts on human societies include damage to coastal communities from sea level rise (Solomon et al., 2007), disruptions in food production due to increased heat stress and drought frequency (Easterling et al., 2007), and more damage from extreme weather events like hurricanes (Emanuel et al., 2008). Impacts on natural ecosystems include extinction
threats to many species resulting from habitat loss (Thomas et al., 2004), expanded ranges for invasive species and ecosystems pests (Fischlin et al., 2007), destruction of coastal ecosystems by extreme weather events and sea level rise (Scavia et al., 2002), and life-threatening changes in water pH and chemistry in aquatic ecosystems (Doney et al., 2009).

The magnitude of these impacts have spurred large-scale international efforts to find climate change mitigation and adaptation solutions, including the meetings leading to the Kyoto Protocol in 1992, the climate conference in Copenhagen in 2009, and numerous efforts by individual nations to constrain greenhouse gas emissions. However, for mitigation and adaptation efforts to be most successful, the budget of atmospheric greenhouse gases under present and future climate conditions must be well constrained. Unfortunately, these estimates and predictions continue to suffer from large uncertainty. Some of this ambiguity results from difficulties in forecasting anthropogenic GHG emission rates over long time scales, which depend on population growth, energy policies, and economic conditions. However, even for well specified emissions scenarios defined by the IPCC, the projected changes in radiative forcing span wide ranges. For example, consider the A1B scenario, which assumes sustained population growth over the course of the next century with the rapid introduction and dissemination of new technologies. The likely range of temperature change associated with the A1B scenario alone is 1.7 to 4.4 °C (Solomon et al., 2007).

The primary source of this uncertainty is the array of dynamic interactions between the earth’s surface and the atmosphere collectively termed "climate change feedbacks". Again, these have been well reviewed elsewhere (Houghton, 2004). Examples include decreased albedo from melting sea ice (which increases oceanic heat uptake), and the cooling impact of increasing cloud cover associated with an accelerated hydrologic cycle.

One of the most significant group of feedbacks, collectively termed "carbon cycle
feedbacks,” includes the complex processes that drive the exchange of CO\(_2\) and CH\(_4\) between the atmosphere and the biosphere. Under present climate regimes, roughly half of annual anthropogenic emissions of CO\(_2\) are cycled through the terrestrial biosphere (Joos, 2003) via photosynthetic carbon assimilation by green vegetation and autotrophic and heterotrophic respiration of carbon. Uncertainties in the balance of net ecosystem assimilation and net ecosystem respiration (i.e. the net ecosystem exchange of CO\(_2\), \(\text{NEE}\)) across the range of global biomes contributes to difficulties in constraining the residual land carbon sink. For example, in the 2007 IPCC assessment report, the residual land sink for CO\(_2\) was estimated to be 2.6 Gt C yr\(^{-1}\), but with an uncertainty range from 0.9 to 4.3 Gt C yr\(^{-1}\) (Denman et al., 2007). The \(\text{NEE}\) flux is sensitive to climate, and the impact of inter-annual climatic variation on ecosystem carbon cycling is believed to be the primary source of inter-annual variability in observed atmospheric CO\(_2\) concentrations (Denman et al., 2007). Ecosystem exchanges of CO\(_2\) and also other GHGs are also sensitive to interactions between the climate and nitrogen cycles. These carbon-nitrogen cycle interactions are particularly complex in wetland and aquatic systems, which are the primary natural source of CH\(_4\) to the atmosphere (Denman et al., 2007).

These present-day uncertainties are magnified in future climate forecasts, when processes such as CO\(_2\) fertilization impacts on photosynthetic assimilation and temperature impacts on respiration must be considered. The forecasts are derived from simulations with Dynamic Global Vegetation Models (DGVMs), which are coupled with General Circulation Models (GCMs) for the climate system. While DGVM outputs are well constrained when applied under certain conditions (Sitch et al., 2008; Zaehle et al., 2005), they may disagree even under current climate conditions (Morales et al., 2005; Siqueira et al., 2006), and they diverge markedly for longer simulation times (Sitch et al., 2008). For example, a recent synthesis involving the most widely used DGVMs showed that projections for the land uptake of carbon at
the end of the 21st century spanned a range greater than the current anthropogenic CO\textsubscript{2} emission rate (i.e. > 7.6 Gt C yr\textsuperscript{-1}) (Purves and Pacala, 2008).

Hence, quantifications of the exchange of GHGs between the biosphere and the atmosphere for both present and future climates is needed for predicting the future climate states that will guide mitigation and adaptation policies. This research need is the primary motivation for the five case studies that comprise the bulk of this dissertation. In the following sections, the complex processes that impact these exchange rates, and past results and knowledge deficits are summarized. New research methods and tools are also discussed. I conclude by outlining the five case studies. I note that the majority of this work concerns the net ecosystem exchange of CO\textsubscript{2} (which remains the GHG associated with the most radiative forcing contributing to long term climate change) and the closely coupled exchange of water vapor. However, in one of the case studies, the exchange of other GHGs (i.e. CH\textsubscript{4} and N\textsubscript{2}O) is also directly quantified.

1.1 Trace gas exchange between the biosphere and the atmosphere - governing processes and complexities

To begin to understand the complexities governing the biosphere-atmosphere exchanges of GHGs, consider the processes governing fluxes of CO\textsubscript{2} and H\textsubscript{2}O (Figure 1.1). Part of the challenge in constraining the atmospheric budget of CO\textsubscript{2} relates directly to the many spatial and temporal scales involved in terrestrial carbon and water cycling. At the cellular and microbial level over short time scales, the rates of carbon exchange in both plants and soils are primarily governed by enzyme kinetics that are sensitive to intrinsic metabolic parameters such as the maximum carboxylation capacity of Rubisco. While these processes are fairly well studied, interesting questions remain, particularly in the context of a rapidly changing climate. For example, the extent to which Rubisco is down-regulated in plants growing in elevated
Figure 1.1: The drivers of the biosphere-atmosphere exchange of CO$_2$ and H$_2$O, and the temporal and spatial scales on which they act. The scales investigated in each chapter of this dissertation are also shown.

CO$_2$ is an active area of research (Rogers and Ellsworth, 2002).

At the leaf and patch scale, the familiar processes of photosynthesis, transpiration ($T_r$), and respiration emerge. While photosynthesis is primarily an aboveground process, transpiration and respiration occurs throughout the vertical gradient of an ecosystem. At these scales, the stomatal conductance to water vapor ($G_s$), which varies with meteorological drivers and soil-plant hydraulics, becomes an important control on rates of carbon assimilation and transpiration. Leaf morphology, and in particularly leaf thickness, is also an important variable governing the diffusion of CO$_2$ to the photosynthetic apparti within leaf cells (Taiz and Zeiger, 2002). In the soil, patch-scale respiration is influenced by temperature and the availability of soil water (Davidson et al., 1998), which can be influenced by microtopography and
small-scale heterogeneities in soil nutrient availability.

Upscaling to the ecosystem scale fluxes like NEE, net primary productivity (NPP) and evapotranspiration (ET) requires consideration of the leaf area present in the canopy (which is commonly measured as leaf area index, or LAI) and the vertical distribution of light within the canopy (as photosynthesis and transpiration are strongly dependent on available radiation). The volume of aboveground and belowground woody biomass is also important for determining NPP, as is the vertical distribution of leaf nitrogen (Wilson et al., 2000). At this scale, the evaporation of water directly from plant elements like leaves and bark or from the soil (E) is an important component of ET, which is the total flux of water vapor per unit area (i.e. \( ET = T + E \), where \( T \) is transpiration per unit ground area).

As the ecosystem-scale fluxes are observed or forecasted over long time scales (i.e. hours to years), variation in NEE and ET are influenced by even more drivers and processes. For example, light availability, temperature, vapor pressure deficit (D) and other meteorological variables known to control photosynthesis and evapotranspiration are driven by weather patterns that may act over times scales of minutes to days (in the case of frontal systems). Seasonal changes in LAI are also a strong control on both processes. The partitioning of carbohydrates to various aboveground and belowground biomass pools (allocation) occurs over time scales of weeks to years and can directly affect the number of leaves in a canopy, increments in aboveground and belowground woody biomass, the efficiency with which plant roots may extract nutrients from the soil matrix, as well as the availability of carbohydrates for heterotrophic respiration in the soil system. The availability of nutrients such as nitrogen and phosphorous to both aboveground plants and soil microbes can also explain flux variability at these temporal scales. Nutrient availability is governed by biogeochemical processes which include the introduction of labile nutrients into the soil system through atmospheric deposition or microbial decomposition of dead
biomass, and the removal of these nutrients through rainwater runoff and occlusion via rock weathering (Schlesinger, 1997).

At long temporal scales the influence of exogenous disturbance events is also critical. These disturbances include forest fires, harvesting, and biomass destruction from insect outbreaks and extreme weather events like ice storms and hurricanes. Until recently, most DGVM’s did not consider disturbance events in simulations and even now, the formulation of these processes needs improvement (Fischlin et al., 2007). This is surprising given that few process can alter the carbon balance of an ecosystem as quickly and completely as the drastic reduction of aboveground biomass common with disturbance events.

Spanning another order of magnitude in time and space is a step necessary to reach the scales associated with long-term climate and land use change. In the spatial dimension, aggregation to these very coarse spatial scales incorporates the full range of global biomes, which themselves span wide ranges of climate conditions and plant species.

It should be noted that many of the interactions shown in Figure 1.1 are dynamic and reciprocal. For example, consider the link between climate change and biophysics. Long-term climate change can impact photosynthesis and respiration, most directly through increased temperature (which drives enzyme kinetics) and the increased atmospheric availability of CO₂ (which drives photosynthesis). At the same time, the net flux of CO₂ and other GHGs can also directly influence climate by either increasing or decreasing the pool of heat trapping gases in the atmosphere. Another example is the link between biophysics and allocation. The availability of carbohydrates available for partitioning to various biomass pools depends strongly on the rate of carbohydrate production via photosynthesis; photosynthesis at the canopy scale, in turn, is largely controlled by LAI, which itself is governed by patterns of allocation. Further, while the bulk of this discussion has so far been focused
on interactions between biophysics and other governing processes, interactions between these governing processes themselves cannot be ignored. For example, the incidence of forest fire and insect outbreaks is largely influenced by meteorology, which in turn is controlled by long term climate states. As another example, species composition change via succession is directly linked to the frequency and magnitude of disturbance events, which remove certain species from the ecosystem and alter the light regime within canopies in ways that favor the growth of other vegetation types.

Finally, while analogous figures and discussions for the processes governing fluxes of CH$_4$ and N$_2$O are not presented here, they are well described elsewhere (Schlesinger, 1997). Sufficient is to say that these processes are equally complex. The challenges and uncertainties associated with constraining the global fluxes of CH$_4$ and N$_2$O are comparable if not greater for other GHGs, which historically have not been as well studied as CO$_2$ (Soussana et al., 2007; Miyata et al., 2000; Rinne et al., 2007), at least at the ecosystem scale. Indeed, measured emissions rates of CH$_4$ and N$_2$O from aquatic ecosystems are highly uncertain, with estimates from different ecosystems that vary by several orders of magnitude (Table 1.1).

1.2 The need for models, and a summary of successes and challenges

Even if all of the processes and interactions that govern the biosphere-atmosphere exchange of GHG were known and quantified, direct scaling from cellular processes to global processes would not be possible given the limits imposed by computer processing speed and memory. Further, direct observation of these fluxes, and the processes that govern them, is not possible for the many distinct terrestrial ecosystems, which means assumptions must be made about the similarity (or dissimilarity) in functioning among ecosystem types. And naturally, future climate states, disturbance regimes, and land cover patterns can not be known with certainty in the present. Therefore, simplified mathematical models are a critical research tool for any scien-
Table 1.1: A comparison of greenhouse gas flux estimates measured recently (i.e. within the last 5 years) for a range of aquatic ecosystems. Units are mg CO$_2$-C m$^{-2}$ h$^{-1}$, mg CH$_4$-C m$^{-2}$ h$^{-1}$, and mg N$_2$O-N m$^{-2}$ h$^{-1}$ for the respective gases. The data represent the range of published values from the season or month associated with the highest fluxes (typically the growing season). 'EC' refers to eddy covariance, 'C' refers to chamber measurements, 'M' refers to mesocosm experiments, and 'FG' refers to flux gradient techniques.

<table>
<thead>
<tr>
<th>Ecosystem</th>
<th>Lat</th>
<th>Lon</th>
<th>Method</th>
<th>Flux</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO$_2$</td>
<td>39° 17' N</td>
<td>122° 11' W</td>
<td>EC</td>
<td>-1728</td>
<td>McMillan et al. (2007)</td>
</tr>
<tr>
<td>estuarine wetland</td>
<td>31° 25' N</td>
<td>121 50' E</td>
<td>EC</td>
<td>-1329 to -481</td>
<td>Alberto et al. (2009)</td>
</tr>
<tr>
<td>flooded rice field</td>
<td>14° 10' N</td>
<td>120 14' E</td>
<td>EC</td>
<td>-709 to -348</td>
<td>Guo et al. (2009)</td>
</tr>
<tr>
<td>tussock tundra</td>
<td>68° 4' N</td>
<td>161° 30' W</td>
<td>C</td>
<td>-381 to -75</td>
<td>Corradi et al. (2005)</td>
</tr>
<tr>
<td>constructed wetland</td>
<td>40° N</td>
<td>83° 1' W</td>
<td>M</td>
<td>-336 to -195</td>
<td>Altor and Mitsch (2008)</td>
</tr>
<tr>
<td>alpine wetland</td>
<td>35° 55' N</td>
<td>101° 21' E</td>
<td>C</td>
<td>-202 to -31</td>
<td>Hirota et al. (2006)</td>
</tr>
<tr>
<td>flooded marsh</td>
<td>49° 20' N</td>
<td>14° 46' W</td>
<td>EC</td>
<td>-197 to -130</td>
<td>Dusek et al. (2009)</td>
</tr>
<tr>
<td>CH$_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>boreal upland sites</td>
<td>55° 55' N</td>
<td>98° 25' W</td>
<td>C</td>
<td>-0.02 to -0.04</td>
<td>Bubier et al. (2005)</td>
</tr>
<tr>
<td>constructed wetland</td>
<td>40° N</td>
<td>83° 1' W</td>
<td>M</td>
<td>0.2 to 1.3</td>
<td>Altor and Mitsch (2008)</td>
</tr>
<tr>
<td>constructed wetland</td>
<td>40° N</td>
<td>83° 1' W</td>
<td>C</td>
<td>1.5 to 3.3</td>
<td>Altor and Mitsch (2006)</td>
</tr>
<tr>
<td>boreal wetland</td>
<td>55° 55' N</td>
<td>98° 25' W</td>
<td>C</td>
<td>1.0 to 7.7</td>
<td>Bubier et al. (2005)</td>
</tr>
<tr>
<td>wetland</td>
<td>46° 58' N</td>
<td>98° 52' W</td>
<td>C</td>
<td>0 to 10.3</td>
<td>Gleason et al. (2009)</td>
</tr>
<tr>
<td>tussock tundra</td>
<td>68° 4' N</td>
<td>161° 30' W</td>
<td>C</td>
<td>3.4 to 10.4</td>
<td>Corradi et al. (2005)</td>
</tr>
<tr>
<td>peat meadow</td>
<td>52° 8' N</td>
<td>5° 3' E</td>
<td>C</td>
<td>11.6 to 11.8</td>
<td>Hendriks et al. (2007)</td>
</tr>
<tr>
<td>treatment wetland</td>
<td>56° 9' N</td>
<td>13° 46' E</td>
<td>C</td>
<td>7.0 to 15.3</td>
<td>Strom et al. (2007)</td>
</tr>
<tr>
<td>treatment wetland</td>
<td>65° 45' N</td>
<td>25° 58' E</td>
<td>C</td>
<td>1.1 to 17.1</td>
<td>Liikanen et al. (2006)</td>
</tr>
<tr>
<td>rice paddy</td>
<td>39° 17' N</td>
<td>122° 11' W</td>
<td>FG</td>
<td>8.6 to 24.4</td>
<td>McMillan et al. (2007)</td>
</tr>
<tr>
<td>alder stand</td>
<td>58° 13' N</td>
<td>26° 47' E</td>
<td>C</td>
<td>1.1 to 27.4</td>
<td>Teiter and Mander (2005)</td>
</tr>
<tr>
<td>rice paddy</td>
<td>22° 12' N</td>
<td>88° 22' E</td>
<td>C</td>
<td>5.2 to 46.7</td>
<td>Mukherjee and Sarkar (2008)</td>
</tr>
<tr>
<td>swamp</td>
<td>29° 48' N</td>
<td>90° 7' W</td>
<td>C</td>
<td>0 to 52</td>
<td>Yu et al. (2008)</td>
</tr>
<tr>
<td>treatment wetland</td>
<td>59° 38' N</td>
<td>11° 19' E</td>
<td>C</td>
<td>-1.2 to 79</td>
<td>Sovik and Klove (2007)</td>
</tr>
<tr>
<td>meadows</td>
<td>34° 44' N</td>
<td>92° 53' E</td>
<td>C</td>
<td>-9.1 to 112</td>
<td>Wang et al. (2010)</td>
</tr>
<tr>
<td>N$_2$O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment wetland</td>
<td>65° 45' N</td>
<td>25° 58' E</td>
<td>C</td>
<td>.01 to .02</td>
<td>Liikanen et al. (2006)</td>
</tr>
<tr>
<td>restored wetland</td>
<td>46° 54' N</td>
<td>98° 42' W</td>
<td>C</td>
<td>.003 to .06</td>
<td>Gleason et al. (2009)</td>
</tr>
<tr>
<td>ponds</td>
<td>50° 12' N</td>
<td>106° 5' W</td>
<td>C</td>
<td>-0.05 to 0.27</td>
<td>Pennock et al. (2010)</td>
</tr>
<tr>
<td>treatment wetland</td>
<td>56° 9' N</td>
<td>13° 46' E</td>
<td>C</td>
<td>-0.2 to 1.2</td>
<td>Strom et al. (2007)</td>
</tr>
<tr>
<td>alder stand</td>
<td>58° 13' N</td>
<td>26° 47' E</td>
<td>C</td>
<td>.06 to 2.0</td>
<td>Teiter and Mander (2005)</td>
</tr>
<tr>
<td>treatment wetland</td>
<td>59° 38' N</td>
<td>11° 19' E</td>
<td>C</td>
<td>-0.02 to 4.6</td>
<td>Sovik and Klove (2007)</td>
</tr>
<tr>
<td>swamp</td>
<td>29° 48' N</td>
<td>90° 7' W</td>
<td>C</td>
<td>-.13 to 8.3</td>
<td>Yu et al. (2008)</td>
</tr>
</tbody>
</table>

The current suite of DGVMs and their stand-scale analogues, via these assumptions and simplifications, attempt to strike a balance between a physiologically sound...
<table>
<thead>
<tr>
<th>Model</th>
<th>Biome- BGC</th>
<th>LPJ</th>
<th>Orchidee</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Step</td>
<td>Daily</td>
<td>Day / Year</td>
<td>Day / Year</td>
</tr>
<tr>
<td>PFTs</td>
<td>Deciduous BL forests</td>
<td>Tropical BL evergreen forests</td>
<td>Tropical BL evergreen forests</td>
</tr>
<tr>
<td></td>
<td>Deciduous NL forests</td>
<td>Tropical BL raingreen forests</td>
<td>Tropical BL raingreen forests</td>
</tr>
<tr>
<td></td>
<td>Evergreen BL forests</td>
<td>Temperate NL evergreen forests</td>
<td>Temperate NL evergreen forests</td>
</tr>
<tr>
<td></td>
<td>Evergreen NL forests</td>
<td>Temperate BL evergreen forests</td>
<td>Temperate BL evergreen forests</td>
</tr>
<tr>
<td></td>
<td>C3 grasses</td>
<td>Temperate BL summergreen forests</td>
<td>Temperate BL summergreen forests</td>
</tr>
<tr>
<td></td>
<td>C4 grasses</td>
<td>Boreal NL evergreen forests</td>
<td>Boreal NL evergreen forests</td>
</tr>
<tr>
<td></td>
<td>Evergreen shrubs</td>
<td>Boreal BL summergreen forests</td>
<td>Boreal BL summergreen forests</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperate herbaceous</td>
<td>C3 grasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tropical herbaceous</td>
<td>C4 grasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C3 agricultural grasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C4 agricultural grasses</td>
</tr>
<tr>
<td>Parameters</td>
<td>42</td>
<td>15</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 1.2: Key features of select Dynamic Global Vegetation Models. *PFT* refers to plant functional types. 'BL' refers to broadleaf, and 'NL' refers to needleleaf. The number of parameters reflects the number of model parameters that must be specified on the basis of *PFT*. The models have other globally specified parameters. The LPJ and ORCHIDEE simulate fast acting processes such as carbon assimilation on a daily time step, and longer processes like allocation and turnover on an annual time step.

representation of governing processes and parsimony in terms of model parameters and computational cost. Most divide the global biomes into plant functional types (*PFTs*), which number between 3 and 20 (Ostle et al., 2009) depending on the model (Table 1.2). A number of model parameters are then assigned on the basis of *PFT* (Sellers et al., 1996; Wang et al., 2007; Zaehle et al., 2005). These parameters primarily describe the impact of meteorology on *NEE* and *ET*, the partitioning of assimilated carbohydrates to various biomass pools, and the rate at which carbon is respired, both autotrophically and heterotrophically. Some models also include explicit formulations of coupled carbon-nitrogen interactions (Xu-Ri and Prentice, 2008; Ostle et al., 2009), disturbance events (Thonicke et al., 2001), soil water redistribution (Siqueira et al., 2008), and other important processes.

These models represent the state-of-the-art in terms of our understanding of plant biosphysics and the interactions shown in Figure 1.1; hence, their formulation
is largely governed by the aspects of this field of research that are well understood. Opportunities for improvements to these models are largely linked to ongoing research into critical knowledge deficits. Some of these “knowns” and “unknowns” are summarized in the following subsections.

1.2.1 Biosphere-atmosphere interactions - some research and modeling successes

Biophysical research has arguably been active for centuries (Katul and Novick, 2007) and has been explored in earnest for decades. At this time, a wide body of literature describing the exchange of trace gases between the biosphere and atmosphere has developed. As illustration, a Web of Science search for ”photosynthesis” returns over 47,000 citations. A similar search for ”transpiration” returns nearly 13,000 citations. With this collective body of work, some important research questions that govern the framework of DGVMs have been successfully addressed. A few pertinent examples are given below.

- The cycles that govern photosynthesis and respiration at the cellular and leaf scales (i.e. the Calvin and citric acid cycles) are well understood (Taiz and Zeiger, 2002); more importantly, perhaps, they are also well described by models. The Farquhar model for photosynthesis (Farquhar et al., 1980), which forms the backbone of many DGVMs, can perform very well when it is appropriately parameterized (Le Roux et al., 1999; Bernacchi et al., 2001; Novick et al., 2004). The link between assimilation and transpiration is also well understood and well represented in models (Katul et al., 2000; Ball et al., 1987).

- The impact of meteorology on carbon assimilation, respiration, transpiration, and associated processes is also generally well understood, at least for short timescales. For example, the relationship between carbon assimilation rates and light is well known at both the leaf and canopy scales (Taiz and Zeiger,
2002; Gilmanov et al., 2003). Models also exist that successfully describe the variation in stomatal conductance with meteorological conditions (Katul et al., 2000), and some are generically applicable across species Oren et al. (1999b). The impact of soil moisture stress imposed by drought conditions on plant function is also well known (Porporato et al., 2001; Reichstein et al., 2007).

• At the stand scale, the distribution of annual NEE for current climate conditions across a wide range of biomes is at least roughly characterized, due in large part to the recent development of a world-wide network of flux towers that employ the eddy covariance methodology (EC) to directly estimated stand-scale fluxes (Baldocchi, 2008). However, questions about how this distribution will shift in future climate persist.

• The upper limit of elevated CO$_2$ enhancements to vegetation growth in future climates is also constrained by past research. Specifically, results from a collection of Free-Air Carbon Enrichment (FACE) studies show that the enhancement effect of elevated CO$_2$ on plant grown is largely limited by nutrient availability (de Graaff et al., 2006; Oren et al., 2001a). Research continues, however, to more precisely quantify the magnitude of this effect.

1.2.2 Biosphere-atmosphere interactions - some remaining challenges

Despite these successes, much uncertainty in model simulations remains (Purves and Pacala, 2008). Some active areas of research aimed at reducing this uncertainty are listed here:

• The influence of temperature and other factors on the emissions of GHGs from boreal peatlands and wetland ecosystems (Denman et al., 2007; Davidson and Janssens, 2006) is still uncertain. Because much of the soil in cold-climate ecosystems is permanently or seasonally frozen, these soils store a dispropor-
tionally large amount of carbon which could be rapidly respired to CO$_2$ or CH$_4$
by soil microbes in warmer climates. The magnitude of methane and nitrous oxide emissions from wetland ecosystems is also poorly constrained (Table 1.1) and spans several orders of magnitude across sites.

- Some emergent processes related to upscaling from the leaf to stand scales are still poorly understood or poorly represented in models. Most importantly, the distribution of light within canopies and the role of diffuse versus direct radiation in carbon assimilation remain important research topics (Baldocchi, 2008).

- The proper specification of the parameters of DGVMs is an active area of research (Wang et al., 2007; Tang and Zhuang, 2008; Friend et al., 2007). Uncertainties in the parameters of terrestrial ecosystem models can inject as much variability in flux estimates as inter-model differences (Knorr and Kattge, 2005).

- The frequency and severity of disturbance events in future climates, as well as the best ways to incorporate these events into model simulations, is critical to constraining future GHG budgets (Fischlin et al., 2007). For example, a recent study suggested that increased incidence of insect outbreaks could cause the continental-scale NEE from Canadian forests to switch from a sink to a source of CO$_2$ in the near future (Kurz et al., 2008).

- There is a need for better characterization of biophysical processes over heterogeneous ecosystems and ecosystems situated in complex terrain. The dearth of research from these types of ecosystems is due in part to the preferential selection of homogeneous, mono-specific study sites and the difficult in applying some measurements techniques in complex settings (Massman and Lee,
2002; Aubinet, 2008). In 2006, a special workshops focused on the challenge of interpreting micro-meteorological measurements from complex terrain was organized in Boulder by the integrated Ecosystem-Atmosphere Processes Study (iLEAPS) organization, resulting in a special issue of *Ecological Applications* devoted to the topic (Aubinet, 2008). More recently, results from the AD-VEX experiment which featured careful attempts to measure advection fluxes have been published in a special issue of *Agricultural and Forest Meteorology* (Aubinet, 2010).

• To an extent, the magnitude of nocturnal greenhouse gases fluxes is also uncertain (Fisher et al., 2007; Baldocchi, 2003). In part, this knowledge gap is related to the prevalence of stable atmospheric conditions during nocturnal periods, which invalidate many of the assumptions necessary for the application of micrometeorological flux estimation techniques (Baldocchi, 2003; Reichstein et al., 2005).

Addressing all of these research gaps is well beyond the scope of a single dissertation, and is likely well beyond the scope of an entire career. Rather, in this body work, certain elements of these research needs are explored via five case studies with the hopes that these results can be combined with ongoing work by other researchers to better formulate DGVMs and better constrain estimates of biosphere-atmosphere trace gas fluxes now and in the future. The content of this work is also governed by the availability of relatively new tools and methods, which are described in the next section.

1.3 Some relatively new research tools

In the past 10 - 15 years, terrestrial ecosystem research has benefitted from the development and/or rapid proliferation of a number of techniques and methods.
Some of the advances most pertinent to this work are discussed below.

1.3.1 The eddy covariance technique

The eddy covariance methodology is a micro-meteorological technique that can, under certain conditions, directly estimate the net ecosystem exchange of mass and energy over spatial scales ranging from $10^2$ to $100^2$ meters. The methodology relies on high-frequency, above-canopy measurements of temperature or scalar concentration and 3-D wind speed (Figure 1.2). For near-neutral atmospheric conditions, these measurements may be incorporated into a simplified version of the mean scalar continuity equation (Baldocchi, 2003), and provided the flux measurement footprint lies within the study ecosystem (Schmid et al., 2000), surface flux estimates may be readily computed for a sufficiently long averaging period (i.e. half-hourly to hourly). The first long-term eddy covariance towers were constructed in the mid-1990s, and since then the number of operating flux towers has expanded to >400 measurement locations from a broad range of global biomes (Baldocchi, 2008).

This technique suffers from limitations primarily imposed by atmospheric conditions that invalidate the simplifying assumptions (the details of which are reserved for subsequent chapters of this text), resulting in large and frequent gaps in the data records. Nonetheless, annual flux estimates are derived using a range of models to gapfill the missing data (Novick et al., 2009b; Moffat et al., 2007). They provide a useful if uncertain constraint on "bottom-up" estimates of net ecosystem exchange which rely on leaf-level gas exchange measurements, allometry, and soil chamber measurements (e.g. Schäfer et al. 2003). However, the most useful application of these datasets may be in the validation and parameterization of DGVMs and other ecosystem models (Friend 2007), and the elucidation of the response of ecosystem fluxes of mass and energy to varying meteorological and edaphic conditions (Stoy et al., 2006a; Reichstein et al., 2007; Novick et al., 2004).
Figure 1.2: An eddy covariance system, comprised of a tri-axial sonic anemometer and high-frequency infrared gas analyzer, situated above a loblolly pine canopy in the Duke Forest, North Carolina, U.S.A.

1.3.2 Remotely-sensed vegetation indices

Aerial and satellite imagery is a powerful tool that can reveal much about vegetation cover over the majority of the earth’s surface. These images may be transformed and de-convolved to provide estimates of the amount of green leaf area through vegetation indices like the normalized vegetation difference index (NDVI, Elmore et al. 2000), which is available at a spatial resolution of 30 by 30 m as part of the LandSAT dataset available from the USGS. LAI is also available for much of the earth’s surface, with time-filled and smoothed datasets at a 1 km by 1 km resolution available from MODIS satellite imagery (Gao et al., 2008). Other ecosystem characteristics such as canopy height (Kellndorfer et al., 2004), land use type, and topographical variation are also available as regional or global datasets. Higher-resolution (i.e. < 5 m by 5
m) imagery from satellites like IKONOS and Quikbird is also available at a price for smaller land areas, and may be manually transformed into vegetation indices. Even higher resolution images are available through LiDAR and other airborne remote sensing platforms.

Clearly, datasets that can characterize, even roughly, the vegetation cover over much of the terrestrial land surface are a critical tool for the validation of regional to global scale models. They may also be employed in present-day model simulations to specify, a priori, the amount and type of vegetation in the study area. These datasets are also useful for detecting and tracking the spread of disturbance events like fires and insect outbreaks (Robinson, 1991; Potter et al., 2005).

### 1.3.3 Advanced analytical methods

Due perhaps to rapid advancements in the PC processing speed over the last 10 - 20 years, researchers in the environmental sciences are increasingly relying on high-powered and complex signal processing and statistical techniques that were largely absent from the field in decades past (e.g. Clark 2007). Some relevant examples include principle components analysis (or PCA, Vaughan and Ormerod 2005), which is a mathematical procedure that transforms an array of potentially non-independent driving variables into a smaller set of uncorrelated variables. Often used in remote-sensing analysis and for large datasets, it is a useful tool in biophysical research as many of the relevant driving variable are not independent (i.e. see Figure 1). Advanced statistical techniques such as neural networks (Papale and Valentini, 2003), the Kalman Filter (Alavi et al., 2006), and multiple imputation (Hui et al., 2004) are often applied to eddy covariance datasets to gapfill missing data. Increasingly, Bayesian methods are being applied as optimization tools to determine likely ranges for model parameters by inverting measured data (Ricciuto et al., 2008; Xu et al., 2006; Trudinger et al., 2007; Fox et al., 2009). And finally, it should be noted that
advanced time series analysis tools, such as wavelet decomposition (Stoy et al., 2005) and power spectra (Siqueira et al., 2006; Mahecha et al., 2010), while they are not necessarily new techniques, are nonetheless increasingly utilized to determine the time scales on which both models and measurements are most variable. Again, when they are employed in this body of work, these methods will be described in more detail.

1.4 The case studies

The five chapters of this dissertation are aimed at decreasing some of the uncertainties in quantifying and forecasting the biosphere-atmosphere exchange of GHGs (previously outlined in Section 1.1). The studies rely on a combination of direct observations and modeling applications and are generally conducted at the ecosystem scale. Whenever possible, they incorporate some of the emerging methodologies and techniques outlined in the previous section. They are briefly summarized here.


Nocturnal evapotranspiration ($ET_N$) is often assumed to be negligible in terrestrial ecosystems, reflecting the common assumption that plant stomata close at night to prevent water loss from transpiration. However, recent evidence across a wide range of species and climate conditions suggests that significant transpiration occurs at night (Figure 1.3), frustrating efforts to estimate total annual evapotranspiration ($ET$) from conventional methods such as the eddy-covariance technique. In this chapter, the magnitude and variability of $ET_N$ is explored in multiple years of eddy-covariance measurements from three adjacent ecosystems in the southeastern U.S.: an old grass field, a planted pine forest, and a late-successional hardwood forest. Relationships between $ET_N$ and meteorological driving variables are quantified, and
Figure 1.3: Histogram describing the occurrence of mean nocturnal transpiration as a fraction of mean daytime transpiration rates from over 50 surveyed species representing both field and greenhouse data. In the case where ranges of percentages were given, the average of the minimum and maximum rates was used. The data are described elsewhere (Novick et al., 2009b).

several gapfilling methodologies for missing data, including process-based multiple non-linear regression, marginal distribution sampling, and multiple imputation, are applied. The utility of the gapfilling procedures is assessed by comparing simulated fluxes to reliably measured fluxes using randomly generated gaps in the data records, and by examining annual sums of ET from the different gapfilling techniques.
Chapter 3. The relationship between reference canopy conductance and simplified hydraulic architecture

Together with canopy leaf area, canopy stomatal conductance regulates plant water use and thereby photosynthesis and growth in terrestrial ecosystems. Although stomatal conductance is coordinated with plant hydraulic conductance, governing relationships across species have not yet been formulated at a practical level that can be employed in large-scale models. In this chapter, combinations of published conductance measurements obtained with several methodologies across boreal to tropical climates are used to explore relationships between canopy conductance rates and hydraulic constraints. A parsimonious hydraulic model requiring sapwood-to-leaf area ratio and canopy height that generates acceptable agreement with measurements across a range of biomes is developed. Further, a relationship between reference canopy conductance and canopy height alone is also developed, and may be more immediately applicable in the terrestrial ecosystem models.

Chapter 4. A characterization of the variability in ecosystem-scale model parameters using Bayesian inversion of eddy covariance data

The parameterization of terrestrial ecosystem models, which is often achieved on the basis of plant functional type, may be a significant source of uncertainty in model applications. In this study, I apply Bayesian inversion techniques to eddy covariance records from over 35 diverse ecosystems to determine the parameters of simple models for the net ecosystem exchange of CO$_2$ and canopy stomatal conductance. Variation in these parameters within and across PFTs and with time is assessed. I also explore relationships between the model parameters and climatic and structural indices that are readily available at coarse-spatial scales in an effort to identify prognostic relationships that could reduce parameter uncertainty.
Chapter 5. Elevated atmospheric CO\textsubscript{2} enhances resin production in field-grown pine trees

This chapter explores the resilience of loblolly pine trees to bark beetle outbreaks, which are among the most destructive disturbance agent in Southern pine forests. In epidemic phases, beetle populations are capable of affecting large land areas with economic and ecological implications comparable to disturbance from forest fires (Kurz et al., 2008; Logan et al., 2003). Projected and in some cases observed increases in beetle populations in warmer climates are widely reported and accepted. However, the concomitant adjustment of the resin flow - the primary defense mechanism against bark beetle attack in trees (Negron et al., 2008) - remains largely unexplored for future climate scenarios. In this study, resin flow measurements in the Free-Air Carbon Enrichment (FACE) facility at Duke University are conducted to assess the impact of elevated atmospheric CO\textsubscript{2}, stand position, and nitrogen amendments on resin production. Observed trends are explained by the growth-differentiation balance hypothesis - a theoretical framework that describes resource allocation to defense mechanisms and that is generally applicable to other plant-herbivore systems.

Chapter 6. Effects of water table depth variations on ecosystem-scale fluxes of greenhouse gases in a large restored temperate wetland

Greenhouse gas exchange rates between the atmosphere and wetland ecosystems are governed by complex processes that vary by orders of magnitude in time and space (Table 1.1). Much of this variation is related to small-scale heterogeneities in vegetation type and microbial communities which in turn are largely controlled by the highly variable water table depth and soil moisture conditions in wetland ecosystems. In this study, two atmospheric flux-estimation techniques that can integrate over smaller scale spatial heterogeneities (the eddy covariance methodology
and a variant of the flux-gradient technique) are employed to determine fluxes of CO$_2$, H$_2$O, CH$_4$, and N$_2$O over hydrologically distinct areas within a large restored wetland in the southeastern U.S. The relationship between flux rates and water table depth is assessed, as is the extent to which emission patterns may be explained by thermodynamic theory. A critical evaluation of the flux-gradient technique and recommendations for future studies are also presented.
2

Nocturnal evapotranspiration in eddy covariance records from three co-located ecosystems in the Southeastern U.S.: Implications for annual fluxes

2.1 Introduction

Until recently, it had been commonly assumed that plant stomata close at night, preventing water loss through transpiration during non-photosynthetic periods. However, sap-flux measurements are now showing that nocturnal transpiration is widespread among woody plants and shrubs (Burgess et al., 1998; Dawson et al., 2007; Donovan et al., 1999; Fisher et al., 2007; Hogg and Hurdle, 1997; Marks and Lechowicz, 2007; Oren et al., 1999a; Snyder et al., 2003; Oishi et al., 2008; Ward et al., 2008). Typically, nocturnal transpiration accounts for 10-30% of the total daily transpiration flux (Figure 1.3), although higher rates have been reported in some extreme cases (Snyder et al., 2003). Hence, models and data collection methodologies that do not sufficiently account for nocturnal water loss due to transpiration may systematically underestimate the contribution of total evapotranspiration ($ET$) to ecosystem water

---

1 This work presented in this chapter is based on work previously published in Novick et al. (2009b)
During the past fifteen years, over 400 eddy covariance towers have been installed across the full range of global biomes to measure ecosystem evapotranspiration and trace gas fluxes (Baldocchi, 2003). The eddy covariance methodology can directly measure whole ecosystem fluxes of water vapor and is often used to develop estimates of annual ET (Schäfer et al., 2002; Alavi et al., 2006; Barr et al., 2007; Burba and Verma, 2005; Gholz and Clark, 2002; Gruenwald and Bernhofer, 2007; Kucharik et al., 2006; Stoy et al., 2006a). However, weak turbulence constraints, which most often occur at night, create frequent gaps that may infect well over 50% of eddy covariance records (Falge et al., 2001).

Much effort has been dedicated to developing appropriate and standard methods of gapfilling missing eddy covariance data for the net ecosystem exchange of carbon dioxide (i.e. NEE, (Falge et al., 2001; Moffat et al., 2007; Reichstein et al., 2005; Stoy et al., 2006b). However, relatively little attention has focused on evaluating and standardizing gapfilling methods for evapotranspiration, and in particular, nocturnal evapotranspiration ($ET_N$). While the two gap-filling problems share some similarities, there are a number of important differences: 1) $ET_N$ is often small and is generated in part by foliage, while night-time respiration is a significant component of NEE and is dominated by soil processes; 2) the meteorological variables impacting $ET_N$ are more numerous and dynamic (e.g., vapor pressure deficit, wind speed, net radiation, air temperature) when compared to respiration processes, which are considered to be governed by a smaller number of variables (i.e. soil temperature and soil moisture) that do not vary appreciably during the night (at least when compared to wind speed and vapor pressure); and 3) many variables that influence daytime ET ($ET_D$) like carbon assimilation and light can not explain variation in $ET_N$ (Adams et al., 2004).

In response to these challenges, a variety of approaches have been developed at
individual sites to gapfill for $ET$ and $ET_N$, including:

1. Process based models such as the Penman-Monteith equation (Anthoni et al., 1999; Stoy et al., 2006a) and the Priestly-Taylor equation (Vourlitis et al., 2002; Wilson and Baldocchi, 2000). These models have the advantage of robustly considering the physiological and some of the physical determinants of daytime $ET$, though it appears that the current suite of models does not appropriately accommodate the potential for large $ET_N$ (Fisher et al., 2007).

2. Look-up tables and similar methodologies that rely on ensemble averages of acceptable data collected under similar meteorological conditions (Falge et al., 2001; Hirano et al., 2003; Law et al., 2000; McCaughey et al., 2006).

3. Multiple linear regression of acceptable fluxes on meteorological variables (Berbigier et al., 2001; Greco and Baldocchi, 1996). This approach incorporates process information empirically, though it does not distinguish between meteorological variables that may influence evaporation and transpiration differently.

4. Setting missing nocturnal $ET$ fluxes equal to zero either directly or indirectly through empirical relationships with photosynthetically active radiation (Novick et al., 2004; Wever et al., 2002).

5. More complex statistical techniques including the Kalman filter (Alavi et al., 2006) and multiple imputation (Hui et al., 2004)).

2.2 Goals of the analysis

Here, the patterns and drivers of $ET_N$ are investigated using long-term eddy-covariance records collected over three ecosystems co-located in the Duke Forest near Durham,
N.C. (U.S.A). Because these three sites experience similar climatic and edaphic conditions, but differ in species composition and leaf area dynamics, a cross-site $ET_N$ comparison is used to elucidate, qualitatively, the relative contribution of evaporation and transpiration to $ET_N$. A range of gapfilling methodologies are applied to the datasets from these three sites and then compared on seasonal and annual time scales. I selected five gapfilling methods of varying complexity: 1) a simple filter that sets all missing $ET_N$ data to 0 (hereafter referred to as ZERO), which is treated as a null dataset for comparison; 2) multiple non-linear regression on covariates of observed meteorological variables (PROCESS); 3) a simple linear regression between the relationship of observed daytime and nighttime fluxes (DAYNIGHT); 4) marginal sampling distribution (DISTSAM) after Moffat et al. (2007), which is a variant of the look-up table approach; and 5) the multiple imputation technique (MULTIMP). The primary criterion for evaluating these five methods is their ability to reproduce artificially generated gaps in data records of acceptable $ET_N$ measurements.

2.3 Materials and methods

2.3.1 Site description

Eddy-covariance derived $ET$ data were collected from three adjacent ecosystems in the Duke Forest in Durham, North Carolina ($35^\circ 58' 41''$ N, $79^\circ 05' 39''$ W) between 2002 and 2006. These ecosystems vary in vegetation cover and canopy structure but experience nearly identical climatic and edaphic conditions. They are an old grass field, a loblolly pine plantation, and a deciduous hardwood forest. Detailed characteristics of each of the ecosystems can be found elsewhere (Ellsworth et al., 1995; Novick et al., 2004; Oishi et al., 2008; Palmroth et al., 2005; Stoy et al., 2006b). However, a brief description is provided here for completeness.

The old field was burned in 1979 and is harvested annually during the summer for hay according to local practices. It is dominated by the grass $Festuca arundi-
*nacea* Shreb., with contributions from other C3 and C4 grasses and various herbs. The height of the vegetation reaches 1 m during the middle of the growing season, concurrent with a peak in leaf area index (*LAI*) of about 2 m$^2$ m$^{-2}$. Eddy covariance instrumentation at this site is positioned 2.8 m aboveground.

The pine plantation was established in 1983 following a clear cut and a burn. *Pinus taeda* L. (loblolly pine) seedlings were planted at 2.0 m by 2.4 m spacing and the stand has not been managed after planting. Canopy height increased from 16 m in 2001 to 19 m in 2006. The canopy is comprised primarily of *P. taeda* with some emergent *Liquidambar styraciflua* L. and a diverse sub-canopy. Eddy covariance instrumentation is positioned at 20.2 m above ground on a 22 m tower. Peak *LAI* at this site was nearly 6.0 m$^2$ m$^{-2}$ in 2001, although it was significantly decreased in later years following repeated droughts and a severe ice storm in 2002 (McCarthy et al., 2006b).

The hardwood forest is classified as an uneven-aged (90 to 110-year old) oak (*Quercus*) – hickory (*Carya*) forest with *L. styraciflua* and *Liriodendron tulipifera* L. also contributing to the canopy and a diverse understory (Pataki and Oren, 2003). The forest has not been managed after establishment. Mean canopy height is 25 m, with the upper canopy, broken by many large gaps, reaching over 35 m. Eddy covariance instrumentation is positioned at 39.8 m on a 42 m tower. Peak *LAI* at this site is 6.3 m$^2$ m$^{-2}$.

All three ecosystems have little topographic variation and lie on Enon silt loam, a low fertility Hapludalf typical of the Southeastern U.S. Piedmont, with a transition to Iredell gravelly loam toward the hardwood forest (Pataki and Oren, 2003). A clay pan underlies the research sites at *ca.* 35 cm below ground (Oren et al., 1998) and imposes similar constraints on root-water access for all three ecosystems. Long-term mean annual temperature and precipitation are 15.5 °C and 1146 mm, respectively.
2.3.2 Evapotranspiration measurements

Latent heat flux \((LE)\) was measured using eddy covariance systems comprised of triaxial sonic anemometers (CSAT3, Campbell Scientific, Logan, UT, USA) and open-path infrared gas analyzers (IRGA, LI-7500, Li-Cor, Lincoln, NE, USA). Measurements of vertical wind velocity and scalar concentrations of water vapor were collected at 10 Hz and flow statistics were processed in real time into half-hour averages using 23X data loggers (Campbell Scientific) with spurious data filtered as described by Katul et al. (1997). The Webb–Pearman–Leuning correction (Webb et al., 1980) for the effects of air density fluctuations on flux measurements was applied. To ensure that mean streamlines are parallel to the ground surface, horizontal wind speed components were rotated to determine an average horizontal wind speed, and then data for periods during which the inverse tangent of the ratio between average horizontal and vertical wind speed exceeded 15° were filtered after Stoy et al. (2006b). The portion of data points filtered through this procedure is less than 0.1%. Measured latent heat fluxes were converted to \(ET\) fluxes, in units of mm per half hour, using the temperature-dependent latent heat of vaporization for water vapor.

2.3.3 Meteorological measurements

A suite of meteorological measurements was also collected at each site. Half-hourly precipitation was measured using a tipping bucket (TI, Texas Instruments, Austin, TX, USA) in the pine plantation. Net radiation \((R_n)\), air temperature \((T_a)\), and relative humidity \((RH)\) were sampled every second and averaged for half-hour periods in all three sites. \(R_n\) measurements were made with Fritschen-type net radiometers (Q7, REBS, Seattle, WA, USA) through 2003 and with CNR1 net radiometers (Kipp & Zonen, Delft, the Netherlands) thereafter. \(T_a\) and relative humidity were measured with HMP35C Ta/RH (Vaisala, Finland) probes positioned at 2 m in the old field and at two-thirds canopy height in the pine plantation and hardwood forest, and
then used to calculate vapor pressure deficit ($D$). Mean horizontal wind speed ($\bar{u}$) and friction velocity ($u^*$) were measured at each site from the velocity components of the sonic anemometer.

2.3.4 Leaf area index measurements

In the old field, $LAI$ was estimated by measuring $PAR$ transmission with a series of 80 quantum sensors (AccuPAR model PAR-80 Ceptometer, Decagon Instruments, Pullman, WA, USA) to calculate gap fractions, which were inverted to calculate $LAI$ after Cambell & Norman (1998). In the pine plantation, the contribution to $LAI$ from $P.\ taeda$ trees was calculated using needle elongation and litterfall measurements, and $LAI$ of understory hardwood species was calculated using degree-day sums and litterfall measurements (McCarthy et al., 2006a). In the hardwood forest, $LAI$ was determined using LAI-2000 (Li-Cor, Lincoln, NE, USA) measurements adjusted by litterfall measurements (Oishi et al., 2008).

2.3.5 Data pre-processing

$LE$ measurements were subjected to a quality control procedure to remove excessive sensor noise and data collected under very stable conditions. First, spikes in the data record were removed using the algorithm recommended for FLUXNET (Papale et al., 2006). Fluxes were then filtered to remove data collected during stable atmospheric conditions using a $u^*$ filtering method similar to the method described by Reichstein and co-workers (Reichstein et al., 2005). Briefly, flux data were sorted into $T_a$ and $u^*$ classes, and a threshold $u^*$ for three-month periods was determined as the value for which mean flux exceeded 95% of the mean flux at higher values of $u^*$. Flux data were then removed if they were collected during periods in which $u^*$ was less than the mean of the three-month threshold series. Finally, $LE$ fluxes were rejected if the half-hourly flux magnitude was outside the logical expectation window of $-100$ to
700 W m$^{-2}$.

When the measurement of any meteorological variable was unavailable due to equipment failure or other error, a continuous record was obtained by fitting a linear regression between measurements from the sensor of interest and a nearby sensor of the same type.

2.3.6 Artificial gap generation

After filtering the data, artificial gaps were created in the data sets for the purpose of evaluating the prognostic ability of the gapfilling methods. Actual gaps in nocturnal eddy covariance records can range in length from one half-hour to well over 50 consecutive nocturnal hours. Due to the scarcity of acceptable nocturnal data and the size of some of the longer gaps in the data set, artificial gaps cannot be created that fully reproduce the gap statistics derived from the entire record. In this analysis, artificial gaps were created by randomly removing all acceptable nighttime data from roughly 10% of the nights.

Data that remained after the quality control filtering and artificial gap generation procedures (hereafter ”acceptable” flux data) were used to evaluate the magnitude of $ET_N$ and its seasonal trends and driving variables. To assess the impact of gapfilling methodology on long-term flux estimates, the data records were then gapfilled as described in the next section.

2.4 Gapfilling techniques

Five gapfilling techniques for $ET_N$ were selected for evaluation in this analysis. These techniques are summarized in Table 2.1 and described below. Three of the methods (PROCESS, DAYNIGHT, and MULTIMP) require model parameters, which were estimated by inverting the acceptable flux data (i.e. the data that remained after quality control filtering and artificial gap creation) using nonlinear optimization
schemes in MATLAB. Nocturnal gaps were then filled with each of the techniques described below. Daytime gaps were filled separately using the Penman-Monteith equation with estimated aerodynamic, boundary layer, and stomatal conductances as described in Stoy et al. (2006a).

2.4.1 Simple filter (ZERO)

I created a ‘dummy’ gapfilled data set by replacing all missing $ET_N$ data with zero. This method reflects the common assumptions that (1) plant stomata are closed and no transpiration occurs at night, and (2) no significant nocturnal evaporation occurs, which is an assumption in any model that drives surface conductance to near zero at night.

2.4.2 Process-based model (PROCESS)

This technique relies on the parameterization of non-linear relationships between $ET_N$ and meteorological variables. To develop the non-linear relationship, I begin by noting that canopy transpiration ($T$) may be represented as (Montheith and Unsworth, 1990):

$$T = a_V \cdot G_S \cdot D \cdot LAI,$$

(2.1)

where $a_V$ is a constant derived from the latent heat of vaporization, the specific heat capacity of dry air, mean air density, and the psychrometric constant. $G_S$ is the canopy conductance to water vapor per unit leaf area, which is commonly expressed with a Jarvis-type reduction function (Jarvis, 1976) such as:

$$G_S = G_{sref} \cdot f(D) \cdot f(t_p),$$

(2.2)

where $t_p$ is the time (in days) since a major rain event (defined as greater than 2mm of rain in a half hour period) – a generic proxy for soil moisture availability.
The functions \( f(D) \) and \( f(t_p) \) adjust a reference conductance rate \( (G_{Sref}) \) to reflect meteorological conditions. Here I chose a reduction function for \( D \) of the form \( f(D) = 1 + m \log(D) \), which describes well the response of \( G_S \) to \( D \) for a wide range of species (Oren et al., 1999b). The soil moisture reduction function takes the form \( f(t_p) = t_p^{-1/2} \), which describes the soil-regulated drying with time following a wetting event (Campbell and Norman, 1998).

The second component of \( ET \) — evaporation (or \( E \)) — may be expressed as a function of \( D \) (Mihailovic et al., 1995):

\[
E = c \cdot D \cdot \frac{1}{r_{s,a}} f(t_p),
\]

where \( c \) is a constant that incorporates \( a_1 \) and information about the relative humidity at the soil surface. The variable \( r_{s,a} \) is the sum of the surface \( (r_s) \) and aerodynamic \( (r_a) \) soil resistances; the former may be represented as a function of the inverse of \( u^* \) (Ye and Pielke, 1993), and the latter is largely a function of the inverse of \( \bar{u} \) (Mihailovic et al., 1995). Noting that \( u^* \) and \( \bar{u} \) are strongly correlated via a constant soil-surface drag \( (C_d = (u^*/\bar{u})^2) \), the soil resistance can be approximated as a function of \( u^* \) alone (i.e. \( 1/r_{s,a} \propto u^{*2} \), Katul et al. 2004). While \( u^* \) was measured at the top of the canopy, the bulk of evaporation occurs below the canopy at the soil-atmosphere interface when the canopy is dry. Therefore, I estimated \( (u^{*2}_{sub}) \) just above the ground surface using measured vertical leaf area distributions with an analytical second-order closure model (Massman and Weil, 1999). Assuming the same soil moisture reduction function for \( E \) gives:

\[
ET = E + T = D \cdot t_p^{-1/2} \left[ (a + b \log(D)) LAI + cu^{*2}_{sub} \right],
\]

where the new constants are \( a = a_1 G_{Sref} \) and \( b = a_1 m \).

The parameters of Equation 2.4 were determined using a non-linear optimization
technique for monthly blocks of data. Average values of the parameters for each month of the year were determined, and combined with Equation 2.4 to fill missing nocturnal data.

2.4.3 Day-night relationships (DAYNIGHT)

Correlation between nocturnal and daytime transpiration rates has been observed in some species (Cavender-Bares et al., 2007) and forms the basis of the DAYNIGHT method, which relies on a simple linear regression model for $ET_N$ as a function of $ET_D$, parameterized for monthly time blocks. For this procedure, analysis was limited to day-night pairs for which three or more acceptable nocturnal half-hour fluxes were available. For months in which less than three such pairs of data were available, the linear relationship was derived from the average slope and intercept obtained during the same months in other years.

2.4.4 Marginal distribution sampling (DISTSAMP)

Marginal distribution sampling is an enhanced look-up table technique (Falge et al., 2001; Moffat et al., 2007; Reichstein et al., 2005) whereby fluxes collected during similar meteorological conditions are sampled within a dynamic window of time centered on the missing data point, and the missing data is replaced with the mean of the sampled fluxes. Similar meteorological conditions were defined as within 0.5 kPA for $D$, 50 W m$^{-2}$ for $R_n$, 0.04 m$^2$ s$^{-2}$ for $u_{sub}^2$, 3 days for $t_p$, and 1 m$^2$ m$^{-2}$ for LAI in a 2-week window centered around the missing flux value. These ranges were selected heuristically to maximize the number of sampled fluxes within the 2-week time window while ensuring similar prevailing meteorological conditions. In the event that less than three data points within this window met the specified criteria, the window size was expanded by two weeks iteratively until an average of three flux rates could be obtained. In cases when three data points meeting the specified criteria
were not available in the entire data record, missing data were linearly interpolated.

2.4.5 Multiple imputation (MULTIMP)

Multiple imputation is a general technique for replacing missing values in large datasets that has recently been used to gapfill eddy covariance records (Hui et al., 2004; Moffat et al., 2007). Briefly, the parameters of a multivariate normal distribution of the acceptable $ET_N$ data and meteorological variables of interest (in this case, $D$, $t_p$, $R_n$, and $LAI$) were estimated using maximum likelihood. $u_{sub}^2$ was not included in the imputation routine because stable periods are typically characterized by low wind speeds and momentum fluxes, and one of the assumptions of the algorithm is independence between the covariates and the likelihood that data is missing.

Missing data were then sampled from this distribution, and the maximum likelihood estimates of the parameters were again estimated using this new dataset. These two steps were iterated until the distribution parameters converge (i.e. the change in the parameters is smaller than an arbitrary threshold set here to $10^{-4}$). At this point, a Markov Chain Monte Carlo algorithm was used to simulate each missing flux value independently using the derived distribution parameters and the corresponding observed meteorological data. Six imputed data sets were generated in this way and pooled into the final gapfilled data set, which includes variance estimates for every gapfilled point.

2.4.6 Estimation of gapfilling error

The accuracy of the gapfilling techniques was assessed with four performance measures: the coefficient of determination ($r^2$), the relative root-mean-square error (RRMSE), the mean absolute error (MAE), and the bias error (BE), where each measure is calculated after Moffat et al. (2007) for each gapfilling technique using observed and simulated data over the artificial gaps. Additionally, frequency distributions of the
Table 2.1: A comparison of the five gapfilling methods used in this analysis. The reference numbers are for: (1) Reichstein et al. (2005), (2) Stoy et al. (2006b), (3) Moffat et al. (2007), and (4) Hui et al. (2004).

<table>
<thead>
<tr>
<th>Method</th>
<th>Ref.</th>
<th>Process Required</th>
<th>Required Variables</th>
<th>Time Window</th>
<th>Computational Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZERO</td>
<td>this study</td>
<td></td>
<td>$ET_N$</td>
<td>N/A</td>
<td>low</td>
</tr>
<tr>
<td>PROCESS</td>
<td>1,2</td>
<td>x</td>
<td>$D, t_p, u_{sub}^2, LAI, ET_N$</td>
<td>monthly</td>
<td>low</td>
</tr>
<tr>
<td>DAYNIGHT</td>
<td>this study</td>
<td>x</td>
<td>$ET_N$</td>
<td>monthly</td>
<td>low</td>
</tr>
<tr>
<td>DISTSAMP</td>
<td>1,3</td>
<td></td>
<td>$D, t_p, u_{sub}^2, LAI, ET_N, R_n$</td>
<td>variable</td>
<td>high</td>
</tr>
<tr>
<td>MULTIMP</td>
<td>4</td>
<td></td>
<td>$D, t_p, u_{sub}^2, LAI, ET_N, R_n$</td>
<td>entire record</td>
<td>high</td>
</tr>
</tbody>
</table>

observed and simulated data are compared to better identify biases in the distributional properties.

For the process-based methods relying on a monthly parameterization scheme (PROCESS and DAYNIGHT), the precision of the gapfilling technique was estimated through a non-parametric bootstrap by randomly selecting monthly parameters from the set of all optimized parameters in the corresponding month. The 95% confidence intervals in the annual and growing season $ET$ estimates were calculated from a Monte Carlo simulation with 100 realizations. The precision of the multiple imputation and marginal distribution sampling methods was not explicitly assessed because, when applied to datasets as large as these eddy-covariance records, these statistical techniques produce very narrow confidence intervals through both parametric and non-parametric bootstrap assessments.

2.5 Results

2.5.1 Meteorological and hydrological variability

The study period was marked by one year of severe growing season (Apr. - Sept.) drought in 2002, when growing season precipitation was nearly 200 mm below the study period average. Growing season and annual precipitation were also well below
the long-term mean in 2005, though cooler than average temperatures and relatively normal precipitation in the early growing season prevented drought status from becoming as severe as in 2002. The sites experienced an exceptionally wet year in 2003, with annual precipitation exceeding the long-term average by 18%. Accordingly, mean annual temperature was nearly 1°C lower in this year when compared to the entire study period. Growing season precipitation rates were closest to normal in 2004 and 2006. The latter year was one of the hottest on record for the study area, and average annual $D$ was approximately 40% higher in 2006 than in 2004.

Precipitation across the three ecosystems is identical, and they do not experience marked inter-site variations in $T_a$ or $D$ (Stoy et al., 2006b). Soil moisture tends to be highest in the hardwood forest in the wintertime when leaf area is largely absent and transpiration drops to nearly zero; soil moisture content is similar across sites during the growing season (Palmroth et al., 2005). Within-canopy wind speed and momentum fluxes vary considerably across the sites (Stoy et al., 2006b).

2.5.2 Data availability

Approximately 6.5% of the ET flux data records were lost due to instrument failures from power outages, sensor malfunctions, and downtime during calibration of the gas analyzers (Table 2.2). The majority of instrument down time occurred during daytime hours when the bulk of maintenance work and calibration was performed. Approximately 9% of data were rejected by the de-spiking routine in all sites, with an additional 38%, 41%, and 42% of data rejected due to filtering across the three sites, the majority of which was collected at night. A small percentage of the data (<3%) was rejected for falling outside of the range $100 \text{ W m}^{-2} < LE < 700 \text{ W m}^{-2}$.

After these filters were applied, 41%, 39% and 45% of the original datasets remained in the old field, pine plantation, and hardwood forest, respectively, although the percentage of nighttime data remaining was much less (18% to 25%). Artificial
gap creation resulted in the exclusion of an additional 2.7%, 1.9%, and 2.8% of data in the old field, pine plantation, and hardwood forest, which represents 12.3%, 9.6%, and 13.9%, respectively, of the acceptable nocturnal data remaining after the data quality filters were employed. The average artificial gap length was approximately 8 half-hour periods (or 4 hours), ranging up to 12 hours.

2.5.3 Presence and seasonal patterns of measured $ET_N$

Mean $ET_N$ over the course of the study period was positive for all sites and the distributions of acceptable $ET_N$ data were all positively skewed (Figure 2.1, with skewness values of 0.63, 0.74, and 0.46 in the old field, pine plantation, and hardwood forest, respectively). Occasionally, negative $ET_N$ was observed, reflecting condensation or random measurement error. $ET_N$ was typically largest in the pine plantation during both summer and winter months, following trends for $ET_D$. Average growing season $ET_N$ was similar in the forest sites (ca. 0.02 mm/hr) and considerably lower in the old field (ca. 0.014 mm/hr). During winter months, when the hardwood forest LAI was near zero, $ET_N$ in the old field was somewhat higher than flux rates in the hardwood forest (0.014 and 0.012 mm/hr, respectively).

Annually, average acceptable $ET_N$ was 8.0%, 9.1%, and 8.0% of $ET_D$ in the old field, pine plantation, and hardwood forest, respectively. The fraction decreased

<table>
<thead>
<tr>
<th></th>
<th>OF Day</th>
<th>OF Night</th>
<th>PP Day</th>
<th>PP Night</th>
<th>HW Day</th>
<th>HW Night</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Lost to power outage</td>
<td>10</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>% Lost to despiking</td>
<td>4</td>
<td>14</td>
<td>4</td>
<td>15</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>% Lost to $u^*$ filtering</td>
<td>20</td>
<td>55</td>
<td>24</td>
<td>58</td>
<td>27</td>
<td>56</td>
</tr>
<tr>
<td>% Lost to thresholding</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>~0</td>
<td>1</td>
<td>~0</td>
</tr>
<tr>
<td>% Lost to artificial gap generation</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total % lost</td>
<td>37</td>
<td>81</td>
<td>40</td>
<td>82</td>
<td>35</td>
<td>75</td>
</tr>
</tbody>
</table>

Table 2.2: The percentage of 30-minute $ET$ flux data removed or lost over the entire study period after progressively employing four data filters and the artificial gap generation procedure.
Figure 2.1: Histograms of the acceptable nocturnal evapotranspiration ($ET_N$) flux data are shown for (a) the old field, (b) pine plantation, and (c) hardwood forest. Dotted lines are drawn at 0 for reference. Frequency represents number of half-hour measurements.

slightly during the growing season, averaging 5.9%, 7.0%, and 7.3%.

Mean monthly $ET_N$ was calculated for all months for which at least 48 acceptable half-hour nocturnal flux measurements were available (Figure 2.2). A strong seasonal pattern in mean monthly $ET_N$ is clearly evident in the hardwood forest, but not in the pine plantation and old field. The hardwood forest experiences large seasonal variations in $LAI$, with a difference of over 5 m$^2$ m$^{-2}$ between peak summertime and wintertime $LAI$. By contrast, $LAI$ in the pine plantation and old field differ by a factor of two or less between summer and wintertime; furthermore, $LAI$ in the
Figure 2.2: The seasonal course for mean monthly $ET_N$ at (a) the old field, (b) pine plantation, and (c) hardwood forest. The squares show the mean flux, and the lines show the standard error of the mean.

Old field experiences rapid fluctuations in summer months as the field is mowed and harvested for hay once or twice each growing season.

In the forested ecosystems, mean annual measured $ET_N$ was lowest in 2002, the year of the extreme drought. Mean fluxes in the old field and pine plantation were high in 2003, which was the exceptionally wet year, and high in 2006 in all sites, which was an abnormally warm year with near-normal precipitation.
2.5.4  Relationship between measured \( ET_N \) and meteorological variables

Half-hourly \( ET_N \) was strongly and nearly linearly related to half-hourly \( u_{sub}^2 \) in the old field and pine plantation, but not the hardwood forest (Figure 2.3 a-c). In the hardwood forest, the largest \( ET_N \) occurred when \( u_{sub}^2 \) was very low, which is characteristic of the growing season (i.e. when high LAI in hardwood forest creates a large momentum sink). At all three sites, \( ET_N \) was strongly related to \( D \) (Figure 2.3 d-f) and increased sharply during periods of strongly negative \( R_n \) (i.e. \( R_n < -100 \text{ W m}^{-2} \), Figure 2.3 g-i), which corresponds to highly negative nocturnal sensible heat fluxes (data not shown). \( ET_N \) appeared to decrease strongly with \( t_p \) during the first week following a major rain event in the pine plantation and hardwood forest (Figure 2.2.4 k-l), although there were no clear trends in \( ET_N \) with \( t_p \) in any of the three sites over much shorter or much longer time scales.

Monthly parameters for the processed based model of Equation 2.4 were successfully determined through optimization routines for 47, 56, and 57 of the total 60 months of the study period in the old field, pine plantation, and hardwood forest respectively. The parameter \( a \), which represents a base nocturnal transpiration rate, was near zero in all three sites. The parameter \( b \) was typically negative in all three sites \( (b = -0.012, -0.0023, \text{ and } -0.014 \text{ mm kPa}^{-2} \) for the old field, pine plantation, and hardwood forest, respectively), which means that the nocturnal stomatal conductance will increase over the base rate \( a \) when \( D < 1 \text{ kPa} \) (and \( b \ln D > 0 \)), which is typical during nighttime. The parameter \( c \) was positive in all three sites, but highest in the old field and pine plantation, reflecting the relationships with \( u_{sub}^2 \) shown in Figure 2.3.

The processed-based model of Equation 2.4 does not robustly reproduced measured fluxes on an hourly time scale. However, when flux records are averaged over longer time periods, the model explains much of the variation in measured \( ET_N \)
However, the model is biased, overestimating small magnitude fluxes and underestimating large magnitude fluxes, with particularly large bias errors observed in the old field.

2.5.5 Gapfilling results

Among the three sites, all five methods estimated the lowest annual flux in 2002, consistent with observed decreases in measured $ET_N$ and $ET_D$ during this extreme drought year (Table 2.3). Further, all gapfilling methods estimate the largest annual $ET$ in 2006 in all three sites, following observed trends in the measured daytime fluxes in this particularly warm year. The magnitude of estimated annual $ET$ was generally highest in the forested ecosystems regardless of gapfilling.

Due to the paucity of nocturnal $ET$ data, the ZERO gapfilling procedure produced annual and growing season estimates that were just marginally higher than the sums of observed $ET_D$ alone. Therefore, most of the variance in the gapfilled estimates from the other four methods can be attributed to the algorithms, which I focus on in this section and the discussion.

In the old field, the various gapfilling methods produced relatively low estimates for the ratio of annual $ET_N$ to $ET_D$ as compared to the pine plantation, with the exception of PROCESS. On average, the gapfilled estimates of $ET_N$ in the old field represented $\sim 6.9\%$ of the mean annual $ET_D$. Across most years in the old field, PROCESS and DAYNIGHT produced nocturnal flux estimates that were typically 1.5 - 2 times as large as the DISTSAMP and MULTIMP estimates, with the exception of 2002, when all gapfilled $ET_N$ estimates were low (i.e. $<40$ mm/yr).

In the pine plantation, gapfilled annual $ET_N$ estimates were approximately 60% higher than in the old field and, on average, comprised a larger percentage of $ET_D$ ($8.5\%$ vs $6.9\%$ in the pine plantation and old field, respectively). The largest estimate of annual $ET_N$ in the pine stand (113 mm, or 15% of $ET_D$) was produced by the
Figure 2.3: The relationship between $ET_N$ and four meteorological drivers: a-c) $u_{sub}^2$, d-f) $D$, g-i) $R_n$, and j-l) $t_p$. The top row is data from the old field (a, d, g, j), the middle row is data from the pine plantation (b, e, h, k), and the bottom row is data from the hardwood forest (c, f, i, l). The circles represent binned averages for discrete classes of the meteorological variables. The error bars show the standard error of the mean for $ET_N$ within each meteorological class. Trendlines are derived from non-linear regression of the binned averages with Equation 2.4. In the case of the relationship between $ET_N$ and $u_{sub}^2$ in the hardwood forest (c), the trendline is derived using averages associated with high $u_{sub}^2$ only, since lower $u_{sub}^2$ is associated with greatly different vegetation cover (i.e. much higher LAI). The trendlines associated with $t_p$ reflect data collected only within the first six days of a major rain event, as no relationship between $ET_N$ and $t_p$ is evident over longer times. Correlation is given for some of the significant relationships discussed in the text.
PROCESS method in 2004, the year with normal temperature and precipitation regimes. In other years, DAYNIGHT generally produced the highest $ET_N$ estimates. Estimates from MULTIMP method were consistently low for all sites and years.

In the hardwood forest, mean annual $ET_D$ was quite stable, varying by $\sim 90$ mm yr$^{-1}$ over the 5-year study period. The range in mean annual estimated $ET_N$ was also similar, with no annual estimates from any of the gapfilling techniques exceeding 100 mm yr$^{-1}$. DAYNIGHT produced the highest $ET_N$ estimates and PROCESS
### Table 2.3: The annual and growing season (GS, Apr-Sept.) evapotranspiration sums (in mm) for the three sites across the five years of the study period. The Daytime Flux is the sum of the Penman-Monteith gapfilled daytime ET fluxes. The other quantities represent the sum of the observed and gapfilled daytime and nocturnal fluxes. The numbers in parentheses show the 95% confidence interval for the ET estimates as determined by a non-parametric bootstrap analysis for the PROCESS and DAYNIGHT methods. \( ET_N/ET_D \) is the ratio of nocturnal to daytime evapotranspiration.

<table>
<thead>
<tr>
<th>Year</th>
<th>Old Field</th>
<th>Pine Plantation</th>
<th>Hardwood Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual ET</td>
<td>GS ET_N/ET_D</td>
<td>Annual ET</td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daytime Flux</td>
<td>482</td>
<td>338</td>
<td>587</td>
</tr>
<tr>
<td>ZERO</td>
<td>486</td>
<td>0.01</td>
<td>340</td>
</tr>
<tr>
<td>PROCESS</td>
<td>503</td>
<td>0.04</td>
<td>346</td>
</tr>
<tr>
<td>DAYNIGHT</td>
<td>521</td>
<td>0.08</td>
<td>356</td>
</tr>
<tr>
<td>MULTIMP</td>
<td>510</td>
<td>0.06</td>
<td>350</td>
</tr>
<tr>
<td>DISTSAMP</td>
<td>517</td>
<td>0.07</td>
<td>350</td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daytime Flux</td>
<td>680</td>
<td>511</td>
<td>708</td>
</tr>
<tr>
<td>ZERO</td>
<td>687</td>
<td>0.01</td>
<td>514</td>
</tr>
<tr>
<td>PROCESS</td>
<td>728</td>
<td>0.07</td>
<td>537</td>
</tr>
<tr>
<td>DAYNIGHT</td>
<td>732</td>
<td>0.08</td>
<td>532</td>
</tr>
<tr>
<td>MULTIMP</td>
<td>707</td>
<td>0.04</td>
<td>523</td>
</tr>
<tr>
<td>DISTSAMP</td>
<td>726</td>
<td>0.07</td>
<td>528</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daytime Flux</td>
<td>611</td>
<td>455</td>
<td>775</td>
</tr>
<tr>
<td>ZERO</td>
<td>616</td>
<td>0.01</td>
<td>457</td>
</tr>
<tr>
<td>PROCESS</td>
<td>654</td>
<td>0.07</td>
<td>482</td>
</tr>
<tr>
<td>DAYNIGHT</td>
<td>660</td>
<td>0.08</td>
<td>478</td>
</tr>
<tr>
<td>MULTIMP</td>
<td>636</td>
<td>0.04</td>
<td>466</td>
</tr>
<tr>
<td>DISTSAMP</td>
<td>650</td>
<td>0.06</td>
<td>471</td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daytime Flux</td>
<td>579</td>
<td>463</td>
<td>792</td>
</tr>
<tr>
<td>ZERO</td>
<td>584</td>
<td>0.01</td>
<td>466</td>
</tr>
<tr>
<td>PROCESS</td>
<td>631</td>
<td>0.09</td>
<td>492</td>
</tr>
<tr>
<td>DAYNIGHT</td>
<td>628</td>
<td>0.08</td>
<td>479</td>
</tr>
<tr>
<td>MULTIMP</td>
<td>606</td>
<td>0.05</td>
<td>475</td>
</tr>
<tr>
<td>DISTSAMP</td>
<td>614</td>
<td>0.06</td>
<td>471</td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daytime Flux</td>
<td>734</td>
<td>546</td>
<td>842</td>
</tr>
<tr>
<td>ZERO</td>
<td>736</td>
<td>0.00</td>
<td>547</td>
</tr>
<tr>
<td>PROCESS</td>
<td>866</td>
<td>0.18</td>
<td>590</td>
</tr>
<tr>
<td>DAYNIGHT</td>
<td>791</td>
<td>0.08</td>
<td>568</td>
</tr>
<tr>
<td>MULTIMP</td>
<td>759</td>
<td>0.03</td>
<td>557</td>
</tr>
<tr>
<td>DISTSAMP</td>
<td>770</td>
<td>0.05</td>
<td>560</td>
</tr>
</tbody>
</table>
Table 2.4: Statistical performance measures for four of the gapfilling methods described in Table 2.1. $r^2$ is the coefficient of determination. RRMSE is the relative absolute error, MAE is the mean absolute error, and BE is the bias error. All four measures were calculated after Moffat et al. (2007).

produced the lowest estimates each year, with the notable exception of 2006.

The simulated fluxes were compared to those measured during the artificial gap periods via the summary performance measures in Table 2.4 and the frequency distributions in Figure 2.5. The MULTIMP method was generally associated with the best RRMSE, though for all methods this quantity is usually within $\sim 20\%$. Estimates of the MAE are generally between 10-18 W m$^{-2}$. Estimates of the bias error exceed 50% in some cases, and are almost exclusively negative due to a general inability of the gapfilling methods to reproduce the largest observed half-hourly fluxes of $ET_N$.

Frequency distributions of the observed and gapfilled data within the artificially generated gaps were examined to further assess the biases associated with the various gapfilling methods (Figure 2.5). In the old field and hardwood forest, DISTSAMP reproduced the shape of the frequency distribution of observed $ET_N$ best (Figure 2.5a and c). In the pine plantation, PROCESS returned the lowest absolute difference between the measured and modeled distributions (Figure 2.5b), followed closely by DISTSAMP and MULTIMP. Generally, MULTIMP, which assumes a multivariate normal distribution between the $ET$ flux and driving variables, failed to reproduce the observed positive skewness in all three sites. The other methods tend to overestimate the degree of skewness, with the exception of DISTSAMP and DAYNIGHT when employed for the old field.
**Figure 2.5:** Relative frequency distributions of the observed acceptable $ET_N$ and four of the gapfilled series for (a) old field, (b) pine plantation, and (c) hardwood forest.

PROCESS and DAYNIGHT, both relying on monthly parameter estimates, were highly sensitive to these parameters as demonstrated by the wide range of annual and growing season $ET$ estimates produced by the bootstrap algorithm (Table 2.3). In some cases, the lower range of the 95% confidence intervals were close to or even below the ZERO flux estimates (e.g., the old field in 2004), although in general the bootstrap estimates suggest a strongly positive $ET_N$. 
2.6 Discussion

The eddy-covariance methodology does not often produce reliable estimates of evapotranspiration due to the frequent suppression of turbulence at night (Falge et al., 2001; Moffat et al., 2007), and even eddy flux measurements collected during near-neutral conditions are often impacted by dynamic flux footprints and sensor noise to the extent that many physiological processes may be obscured (Oren et al., 2006). Nonetheless, in this study, a significant nocturnal evapotranspiration signal from eddy covariance records was detected at all three study sites (Figures 2.1 and 2.2) that accounted for an average of 8.0%, 9.1%, and 8.0% of the annual daily ET flux in the old field, pine plantation, and hardwood forest respectively. These ratios are at the low end of the range of nocturnal ET flux magnitudes reported in other studies that rely primarily on sap flux or porometry data (see Figure 1.3). The estimates are likely underestimated by 12-20%, which is the underestimation range in nocturnal NEE derived from eddy-covariance (Soegaard et al., 2000; Wilson et al., 2002), which is similar to the underestimation of $ET_N$ in a recent study (Fisher et al., 2007).

In the following subsections, I further discuss errors for, and independent checks on, these estimates. I also discuss the contribution of $T$ and $E$ to $ET_N$, as well as the relationships between these fluxes and meteorological drivers. I then evaluate the accuracy of the gapfilling methodologies employed in this analysis, and conclude by discussing the significance of these findings with respect to ecosystem water budgets and ecophysiological modeling.

2.6.1 Measured $ET_N$ – sources of error

The data quality filtering scheme resulted in the rejection of a large portion of the data records (ca. 60%, Table 2.2), which is at the upper limit of data rejection reported from other sites (Moffat et al., 2007). While the proportion of the data
rejected by these procedures is similar to that derived from another quality control assessment employed at these sites incorporating an atmospheric stability filter in place of $u^*$ filtering (Stoy et al., 2006b; Novick et al., 2004), it is possible that acceptable data instructional to the analysis and gapfilling has been removed while attempting to minimize systematic errors associated with stable boundary conditions.

To address this concern, $MAE$ in the nocturnal $LE$ measurements was calculated by comparing acceptable flux measurements collected during the same time period on consecutive days under similar environmental conditions ($u$ within 1 m s$^{-1}$, $D$ within 0.5 kPa, and $R_n$ within 50 W m$^{-2}$). Flux error estimated using this daily differencing technique has been previously shown to follow a double-exponential (Laplacian) distribution, for which the scale parameter is $\delta$ (Richardson et al., 2006). The derived estimates of $MAE$ (9.19, 12.20, and 8.26 W m$^{-2}$ for the old field, pine plantation and hardwood forest, respectively) are within the range of errors in $LE$ reported for seven other Ameriflux sites (3.4 - 15.6 W m$^{-2}$ for $R_n < 100$ W m$^{-2}$) (Richardson et al., 2006).

In all three sites, $MAE$ is directly related to $u^*$ ($r^2 = 0.7$, 0.84, 0.62 in the old field, pine plantation, and hardwood forest, respectively). Because most data collected during stable boundary conditions is associated with low $u^*$, removing these data points should not result in large increases in the uncertainty of flux estimates. Indeed, the estimated ratio of nocturnal to daytime $ET$ varies by <10% in all three sites when additional artificial gaps are created in the data records to simulate up to 90% "loss" of the data.

However, I cannot entirely discount the possibility that greater data availability would impact flux estimates, especially considering that much more data is missing during the growing season. Because $ET_N$ is highest during the growing season, the greater portion of missing data during summer months is likely to inject a negative bias in flux estimates. Therefore, estimates of the magnitude of $ET_N$ presented in
this study should be considered as conservative minima.

2.6.2 Independent validations

Some independent checks of the nocturnal ET estimates can be found in porometric gas exchange and sap flux measurements conducted at the study sites. In the pine plantation, leaf-level gas exchange measurements have been collected as part of a separate study and used to parameterize the Ball-Berry model for stomatal conductance (Ball et al., 1987; Collatz et al., 1991), which may be expressed as:

\[ G_S = m' \frac{A \cdot RH}{c_a} + b' \]  

(2.5)

where \( A \) is the CO\(_2\) assimilation rate, \( RH \) is relative humidity, \( c_a \) is the ambient CO\(_2\) concentration, and \( m' \) and \( b' \) are constants. The latter constant, which has been measured as 0.015 mol m\(^{-2}\) s\(^{-1}\) in the pine plantation (Lai et al., 2000), represents the stomatal conductance during non-photosynthetic periods (i.e. nighttime). Using this baseline stomatal conductance rate in Eq. 2.1 together with observed \( D \) and LAI generates estimates for annual \( ET_N \) in the pine plantation between 56 and 91 mm/yr over the course of the study period, or between 10-18% of \( ET_D \). These estimates are comparable to those produced with the eddy-covariance gapfilling procedures (Table 2.3).

In the hardwood forest, sap-flow in dominant canopy species was measured using heat dissipation probes as part of a separate study performed concurrent to this investigation (Oishi et al., 2008). The ratio of nocturnal to daytime water flux through the stem determined from sap flow measurements averaged 17±0.2% during the period of 2002 to 2005. This is within the range of ratios reported by other studies (see Figure 1.3), although it is approximately twice the observed ratio of acceptable \( ET_N/ET_D \) from eddy covariance fluxes. The discrepancy between eddy
covariance estimates of $ET_N$ and sap-flux derived estimates of nocturnal $T$ in the hardwood forest may be explained in part by the known systematic underestimation of nocturnal eddy covariance fluxes. Further, a portion of the nocturnal water uptake measured with sap flow probes is used for water recharge in the trees, and would not be sensed by the eddy covariance instrumentation.

2.6.3 Determinants of measured $ET_N$

Measured $ET_N$ was related to $D$ in all three sites, consistent with previous studies (Daley and Phillips, 2006; Dawson et al., 2007; Fisher et al., 2007; Kavanagh et al., 2007; Oren et al., 2001b; Ward et al., 2008). It was strongly correlated with $u_{sub}^{*2}$ in the old field and pine plantation, and in the hardwood forest during periods of low LAI (i.e. higher $u_{sub}^{*2}$). In all three sites, $ET_N$ did not appear to be strongly driven by $t_p$ except within the first seven to 10 days following a rain event. This scale separation may simply reflect the influence of noise on smaller fluxes during dry periods, or suggest the presence of more complex processes such as hydraulic redistribution that alter relationships between $ET$ and $t_p$ (Siqueira et al., 2008). Further, the $t_p^{-1/2}$ relationship applies only during the second-stage of soil evaporation, and does not necessarily reflect the final stages of evaporation.

The dependence of $ET_N$ on $u_{sub}^{*2}$ is not surprising given the results of previous studies (Phillips et al., 2010), but nonetheless warrants further process-based exploration. Specifically, I asked: “Could the direct relationship between $ET_N$ and $u_{sub}^{*2}$ simply reflect varying boundary layer and aerodynamic conductances within the canopy, as opposed to a meteorological control on canopy stomatal conductance?”

To answer this question, I focused on $ET_N$ from the hardwood forest collected during the peak of the growing season (i.e. July), when high leaf area index and low $\bar{u}$ and $u^*$ would promote relatively low boundary layer and aerodynamic conductance. Using this data, I estimated nocturnal leaf boundary layer conductance as
\[ G_{BL} = 255 \sqrt{u/d}, \] where \( d \) is a characteristic leaf dimension calculated after Stoy et al. (2006a). This conductance was estimated as a vertical profile throughout the canopy using wind speed estimates from the 2nd-order closure model of Massman and Weil (1999).

The aerodynamic conductance was estimated according to

\[
G_{aero} = \frac{\kappa u^*}{\ln \frac{z-d}{z_o} + \psi} \quad (2.6)
\]

where \( \kappa \) is the von Karman constant, \( d \) is the zero plane displacement, and \( z_o \) is the roughness height. The stability correction term \( \psi \) was formulated after Kader and Yaglom (1990) as a function of the Obukhov length, which may be estimated from measured sensible heat flux and \( u^* \).

At all positions within the canopy, \( G_{BL} \) is typically \( > 5000 \text{ mmol m}^{-2} \text{ s}^{-1} \) (Figure 2.6a). \( G_{aero} \), estimated as a canopy average, is typically \( > 1000 \text{ mmol m}^{-2} \text{ s}^{-1} \) (Figure 2.6b). As canopy stomatal conductance rates per unit leaf area do not typically exceed \( 1000 \text{ mmol m}^{-2} \text{ s}^{-1} \) even under high-light conditions (Novick et al., 2009a), I conclude that for most of the acceptable \( ET_N \) measurements, stomatal conductance is the limiting conductance to water vapor transfer from the leaf to the atmosphere.

Significant relationships between nocturnal and daytime transpiration rates were observed for at least some months in all three sites, consistent with previous findings (Cavender-Bares et al., 2007). The greatest \( ET_N \) fluxes were observed in the pine plantation (Figure 2.2, Table 2.3), which is dominated by the fast-growing, shade-intolerant \( P. \) taeda. This result is consistent with previous observations that nocturnal transpiration is highest in species with high extension growth rates and low shade tolerance, and may represent a tendency for these species to trade water for enhanced nocturnal carbon export driven by dark respiration (Marks and Lechowicz, 2007).
Figure 2.6: The vertical profile of average nocturnal leaf boundary-layer conductance estimated for the hardwood forest during July is shown in (a). A histogram of the canopy-averaged aerodynamic conductance for the same time periods is shown in (b).

2.6.4 Influence of $T$ and $E$ on $ET_N$

Partitioning the measured above-canopy $ET$ fluxes into its $E$ and $T$ components is challenging even during daylight hours (Roupsard et al., 2006; Stoy et al., 2006a; Williams et al., 2004), and I will not attempt a formal quantification of the contribution of $T$ and $E$ to $ET_N$. Instead, I discuss the relative contribution of these fluxes qualitatively.

In the hardwood forest, perhaps the strongest evidence that a significant portion of the observed $ET_N$ is transpiration comes from the seasonal trends in monthly
averaged $ET_N$ (Figure 2.2). If $ET_N$ in the hardwood forest was dominated by $E$, then $ET_N$ should decline during summer months, when increased leaf area reduced the wind speed and momentum at the canopy floor (i.e. $u_{sub}^2$ is at its lowest), and soil moisture content was substantially lower. Instead, $ET_N$ rates were substantially higher during the growing season (see Figure 2.2c), suggesting a significant contribution from $T$.

No seasonal pattern of monthly $ET_N$ is evident at pine plantation; however, $LAI$ is high at this site year-round, and data from this site were most often missing during summertime, which are both conditions that may obscure seasonal differences in the flux record. In both forested ecosystems, the response of $ET_N$ to the primary meteorological variables $D$ and $u_{sub}^2$ is not as strong as in the old field, suggesting a biological control of $ET_N$ that, with respect to $D$, may reflect reduced stomatal conductance at high $D$.

A priori, I expected that nocturnal $E$ would be highest in the old field due to the low degree of canopy cover, which generally increases soil evaporation (Mihailovic et al., 1995). The greater magnitude of $ET_N$ in the pine plantation relative to the old field during all seasons, and in the hardwood forest relative to the old field during the growing season, suggests $ET_N$ includes significant contributions from $T$ in the pine plantation and hardwood forest. And finally, the magnitude of nocturnal sensible and latent heat fluxes is more similar in the old field than in the forests (data not shown), suggesting a greater contribution to $ET_N$ from $E$ in the old field since nighttime evaporation fluxes are typically induced by large downward sensible heat fluxes (Fisher et al., 2007).

2.6.5 Evaluating gapfilling methods for $ET_N$

In all sites, the DAYNIGHT and DISTSAMP methods consistently ranked highly for at least two of the four performance measures presented in Table 2.4. The coefficient
of determination is quite small for all techniques, which is not surprising given that
the magnitude of $ET_N$ at these sites (.01 to .02 mm hr$^{-1}$ or 9 to 13 W m$^{-2}$) is on the
same order of estimates of the random measurement error in $ET$ during low radiation
periods (3.4 to 15.6 W m$^{-2}$) according to Richardson et al. (2006). Therefore, this
is not viewed as a reliable performance measure.

The $MAE$ associated with DAYNIGHT and DISTSAMP is typically within 2 W
m$^{-2}$ of the $MAE$ estimated for the data from the paired-point technique described in
Section 2.6.1, suggesting that the uncertainty associated with these two methods is
close to the random error in the data. Generally, all gapfilled fluxes have a negative
bias error, due to the inability of the methods to reproduce extremely large fluxes;
however, the bias error associated with DAYNIGHT and DISTSAMP are within
$\sim$20%, as compared to errors that may exceed 50% for the other methods.

Among all sites and years, the MULTIMP flux estimates were consistently low
and the DAYNIGHT flux estimates were consistently high. For MULTIMP, the
apparent underestimation of nocturnal $ET$ flux rates is directly linked to the as-
sumption of multivariate-normality in the joint-distribution between $ET$ and the
driving variables, which contradicts the prominent positive skewness observed in the
data (Figures 2.1 and 2.5). In DAYNIGHT, all missing nighttime data were replaced
with a constant value for each night obtained using relationships developed with the
mean of observed nighttime values. Because data were more likely to be observed at
high wind speeds (when atmosphere is near-neutral conditions), the measurements
used to parameterize these relationships are also high.

In all three sites, the DISTSAMP was one of the top performing gapfilling es-
states in terms of an ability to reproduce the relative frequency distributions of
measured $ET_N$ (Figure 2.5). This method is also not susceptible to large variation
in annual $ET_N$ driven by sensitivity to model parameter estimation, as are both
PROCESS and DAYNIGHT.
Based on these assessments, I conclude that the DISTSAMP method is the best performing gapfilling method examined in this study. However, all the methods, including the DISTSAMP, are negatively biased with respect to the observations, which in turn may be negatively biased with respect to the true fluxes due to under-representative sampling during the growing season and well-known measurement and systematic errors associated with nocturnal eddy covariance measurements (Wilson et al., 2002). Therefore, all gapfilled estimates of $ET_N$ represent conservative minimum annual flux estimates, and more work is necessary to elucidate the drivers of nocturnal evapotranspiration and to modify gapfilling methods accordingly.

2.6.6 Broader implications and conclusions

The evidence for significant nocturnal $T$ challenges the long-held notions that plant stomata close at night and that the primary drivers for ecosystem water use are only important during daylight hours. Nocturnal evapotranspiration, which until recently has been largely ignored in ecosystem water budgets, appears to be a non-trivial factor in site water balances (Dawson et al., 2007). This new evidence invalidates many models and gapfilled eddy-covariance $ET$ estimates (Fisher et al., 2007), necessitating a more robust treatment of nocturnal $ET$ to remove biases in long-term flux estimates.

Little is known about the drivers and functions of nocturnal $ET$ and explicit incorporation into process-based models is needed to set testable hypotheses. Some studies suggest that nocturnal transpiration functions to prevent CO$_2$ accumulation at night (Marks and Lechowicz, 2007), or to enhance nutrient uptake (Scholz et al., 2007), while other suggest that incomplete stomatal closure is the primary determinant of nocturnal water loss (Oren et al., 2001b; Marks and Lechowicz, 2007). Here, the results show that $ET_N$ is detectable and correlates with driving variables despite the noise and numerous gaps infecting eddy covariance datasets as well as the absence
of a clear process-based explanation for the presence of nocturnal water loss. The DISTSAMP method used here was shown as the best for gapfilling eddy covariance ET. Until better process-based models are available, the large differences in annual ET estimates among these methods compel us to recommend checking gapfilled ET against independent estimates of nocturnal water use (i.e. from sap flux, porometry, or isotope data) or, at the least, performing a frequency distribution analysis similar to those performed here and searching for potential biases in the gapfilled records.
The relationship between reference canopy conductance and simplified hydraulic architecture

3.1 Introduction

Canopy stomatal conductance to water vapor ($G_S$) is a primary determinant of ecosystem transpiration rates. Over the past few decades, much attention has been focused on describing the response of $G_S$ to variables that act on fast time scales (e.g. hourly). In comparison, little attention has been paid to processes that may impact canopy conductance on longer time scales (e.g. yearly). Generic relationships that are valid across species have been developed for the fast responses of $G_S$ to $\text{PAR}$ (Granier et al., 2000), $D$ (Oren et al., 1999b), and soil moisture ($\theta$ Granier et al. 2000) and have been implemented in large-scale models. These models typically rely on a reference canopy conductance rate ($G_{Sref}$), defined at a specific environmental state that can vary across applications and adjusted for the fast-acting meteorological variables. These adjustments can be based on multiplicative functions (see Equation 2.2), which in this study are $f_1(D)$, $f_2(PAR)$, and $f_3(\theta)$.

---

1 This work presented in this chapter is based on work previously published in Novick et al. (2009a)
$G_{S_{ref}}$ significantly varies across stands of different age, structure and vegetation type, and changes predictably with measurable features of canopy structure, at least within a species (Schäfer et al., 2000; Delzon et al., 2004; Ryan et al., 2006). However, the current suite of terrestrial ecosystem models do not account for mechanisms that impact $G_{S_{ref}}$ over longer time scales. Some DGVMs and stand-level models assume that the canopy stomatal conductance parameters are ‘static’ for a range of canopy architectural scenarios, while others change the parameters empirically with stand age, or require species-specific allometric relationships that are difficult to implement over large and biologically diverse land areas (Kucharik et al., 2006; Siqueira et al., 2006). Traditionally, these assumptions were necessary given the lack of spatial datasets of elementary hydraulic parameters known to impact $G_S$. Recent advances in LiLight Detection and Ranging (LiDAR) imaging technology now facilitate detailed mapping of key properties of canopy architecture for large land areas (Lefsky et al., 1999, 2002), and elevation datasets from the Shuttle Radar Topography Mission appear capable of producing maps of canopy height ($h$) over most of the global land surface (Kellndorfer et al., 2004).

3.2 Theoretical considerations and hypotheses

3.2.1 Relating transpiration and conductance to hydraulic architecture

The cohesion-tension theory for water transport in trees (Tyree, 1997) has been used to explain the contribution of hydraulic characteristics to variations in $G_S$. Within species, theoretical relationships between canopy stomatal conductance and canopy architecture are often derived by equating the soil-to-leaf water flux to the leaf-level transpiration rate ($T_r$, mmol m$^{-2}$ s$^{-1}$) under steady-state flow conditions (Whitehead et al., 1984; Tyree and Ewers, 1991), yielding:

$$T_r = K(\Psi_{soil} - \Psi_{leaf} - \rho_w gh),$$  \hspace{1cm} (3.1)
where $K$ (mmol m$^{-2}$ s$^{-1}$ MPa$^{-1}$) is the leaf-level hydraulic conductivity from the soil to the leaf, $g$ is gravitational acceleration (m s$^{-2}$), $\rho_w$ is the density of water (kg m$^{-3}$), and $\Psi_{soil} - \Psi_{leaf}$ (MPa) is the soil-to-leaf pressure difference. Noting that $K$ is proportional to the sapwood area and inversely proportional to soil-to-leaf path length (Oren et al., 1999a; Schäfer et al., 2000) yields:

$$T_r = k_s \frac{A_S}{A_L h} (\Psi_{soil} - \Psi_{leaf} - \rho_w gh),$$

(3.2)

where the path length from $\Psi_{soil}$ to $\Psi_{leaf}$ is approximated by $h$, and $k_s$ is the tissue-specific hydraulic conductivity per unit sapwood area (mmol m$^{-2}$ s$^{-1}$ MPa$^{-1}$).

Ecosystem- and coarse-scale carbon cycling models often assume that, at long time scales, leaf boundary layer conductance has negligible influence on total canopy conductance. With this assumption, the stomatal response to changes in hydraulic architecture can be predicted by substituting $G_S$ and the vapor pressure deficit ($D$, MPa) for the transpiration rate in Equation 3.2 (Jarvis, 1976; Whitehead, 1998; Ewers et al., 2000; Addington et al., 2006), yielding:

$$G_S D = k_s \frac{A_S}{A_L h} (\Psi_{soil} - \Psi_{leaf} - \rho_w gh).$$

(3.3)

3.2.2 Separating fast and slow responses

As noted earlier, $G_S$ responds rapidly to changes in $PAR$, $D$, and $\theta$ via the multiplicative functions $f_1(D)$, $f_2(PAR)$, and $f_3(\theta)$. Therefore, to isolate the effects of $A_S/A_L$, $k_s$, $\Psi_{leaf}$, and $h$ on $G_S$ from the effects of rapidly changing variables, a conductance rate at a reference environmental state ($G_{Sref}$) is used. In this analysis, the reference environmental state is characterized by non-limiting light and soil moisture (i.e. $f_2(PAR) = f_3(\theta) = 1$), and a reference $D$ of 1 kPa. Estimates of $G_{Sref}$ may be adjusted to reflect varying environmental conditions to produce a continuous
estimate of \( G_{Sref} \) as per Equation 2.2 with multiplicative functions. In the case of inter-specific application of the Jarvis model, at least one variant of the three functions \( f_1(D) \), \( f_2(PAR) \), and \( f_3(\theta) \) have already been formulated (see Oren et al. 1999 for \( f(D) \), and Granier et al. 2000 for \( f(PAR) \) and \( f(\theta) \)).

When only non-limiting soil moisture states are considered (as specified by the reference environmental state), \( |\Psi_{soil}| \) is typically an order of magnitude less than \( |\Psi_{leaf}| \). Therefore, \( |\Psi_{soil}| \) is neglected in Equation 3.3 relative to \( |\Psi_{leaf}| \), noting that this may introduce a bias on the order of 10-20% in plants with relatively low \( |\Psi_{leaf}| \) (Figure 3.1). With this assumption, \( G_{Sref} \) can be expressed as a function of \( A_S/A_L \), \( k_s \), \( \Psi_{leaf} \), and \( h \) using:

\[
G_{Sref} = k_s \frac{A_S}{A_L h} (\Psi_{leaf} - \rho_w g h).
\] (3.4)

This formulation assumes that canopy height is a proxy for the mean path length from the soil through the rooting zone to the leaf. Conditions in which \( h \) does not represent this path length for water flow are likely to occur in two types of ecosystems: a) canopies with deep rooting relative to the total path length (i.e., mature short stature forests), and b) canopies where complicated vertical branch architecture patterns make \( h \) a poor proxy for the mean path length. In the former scenario, a rooting length of 1 m results in a 5% error in \((k_s/h) \cdot (A_S/A_L)\) for a 10 m canopy (Figure 3.1). Similarly, a rooting length of 2 and 3 m result in errors of ca. 15% and 20%, respectively. In the case of taller canopies, the error introduced by equating \( h \) with the path length decreases with increasing \( h \).

To assess the relative contribution of each of these four variables to inter-specific variation in the reference conductance rates, the observed natural variation in these parameters is considered first. In general, \( \Psi_{leaf} \) is typically around -2 MPa (Delzon et al., 2004; Koch et al., 2004; Meinzer, 2003), although values as high as \( \Psi_{leaf} = -1.0 \)
Figure 3.1: The error introduced by some of the assumptions leading to Equation 5. (a) shows the error in \((k_s/h) \cdot (A_s/A_L) \cdot (\Psi_{\text{leaf}} - \rho_wgh)\) incurred by neglecting root length (RL) in the total path length for a range of assumed root depths. (b) shows the relative error associated with neglecting \(|\Psi_{\text{soil}}|\), which is typically an order of magnitude less than \(|\Psi_{\text{leaf}}|\) for a range of soil water potentials. The dotted lines indicate 10 and 20 % errors.

MPa (\textit{Picea mariana}, Ewers et al. 2005) and \(\Psi_{\text{leaf}} = -1.1\) MPa (\textit{Eucalyptus saligna}, Barnard and Ryan 2003), and as low as \(\Psi_{\text{leaf}} = -3.28\) MPa (tropical species, Sobrado 1986) and even much lower have been reported. The hydraulic conductivity, \(k_s\), varies across species by about an order of magnitude, from < 30 mmol m\(^{-1}\) s\(^{-1}\) Mpa\(^{-1}\) for gymnosperms to > 130 mmol m\(^{-1}\) s\(^{-1}\) Mpa\(^{-1}\) for some evergreen angiosperms (Maherali et al., 2004).

Variations in \(A_s/A_L\) across species are comparable to variations in \(k_s\), ranging
from values as low as 0.7 cm$^2$ m$^{-2}$ for tropical *Eucalyptus saligna* (Barnard and Ryan, 2003) and 0.5 cm$^2$ m$^{-2}$ for boreal species (Ewers et al., 2005) to ratios as high as 13 cm$^2$ m$^{-2}$ for *Pinus palustris* (Addington et al., 2006) and 14 cm$^2$ m$^{-2}$ for *Taxodium distichum* (Oren et al., 2001b). Even greater variations are found over the landscape in $h$, which can range from $< 1$ m to over 100 m.

Therefore, if independence is assumed among all of the driving variables in Equation 3.4, the products $k_s(\Psi_{leaf} - \rho wgh)$ and $A_S/A_L/h$ should both vary by approximately an order of magnitude across species, and each group of variables could explain roughly 50% of the inter-species variation in $G_{Sref}$ if all other assumptions in the model are valid. In actuality, some coordination among these variables is likely. For example, within species, $A_S/A_L$ and $h$ are often tightly correlated (Schäfer et al., 2000; McDowell et al., 2002a; Ewers et al., 2005) and linked by a simple linear relationship:

$$A_S/A_L = \alpha_S h + \beta_S. \quad (3.5)$$

However, $\alpha_S$ can be either positive or negative (McDowell et al., 2002a), and can vary from as low as -0.41 cm$^2$ m$^{-3}$ (*Picea mariana*, Ewers et al. 2005) to as high as 0.21 cm$^2$ m$^{-3}$ (*Pinus sylvestris*, Magnani et al. 2000). Hence, across species, $A_S/A_L$ and $h$ are expected to be less correlated than among stands of the same species. Furthermore, compensating relationships between $\Psi_{leaf}$ and $k_s$ should be considered. Trees growing in dry environments conducive to producing low (i.e., more negative) $\Psi_{leaf}$ produce tissues with lower xylem vulnerability to cavitation accompanied by lower $k_s$ (Sperry et al., 1998; Ewers et al., 2000). Conversely, plants producing tissues with high $k_s$ must maintain higher $\Psi_{leaf}$ to prevent xylem cavitation (Tyree and Sperry, 1988). Thus, a change in $\Psi_{leaf}$ that could have a positive effect on $G_{Sref}$ would probably be accompanied by an opposing change in $k_s$ and vice versa. A
recent review article found that the leaf water potential associated with a 50% loss in hydraulic conductivity explained about 60%

This analysis is focused on the relationship between $G_{Sref}$ and $A_S/A_L/h$ as canopy height is an easily measurable feature of canopy architecture, and sapwood-to-leaf area is far simpler to measure at the stand-scale than $G_{Sref}$. Furthermore, $A_S/A_L$ may be determined \textit{a priori} for some species based on established allometric relationships or LiDAR remote sensing. I hypothesize that, hydraulically, $A_S/A_L$ and $h$ should exert a strong control over $G_{Sref}$, explaining approximately 50% of the variation in reference conductance via:

$$G_{Sref} \propto A_S/A_L/h.$$ \hspace{1cm} (3.6)

Within this framework, results from two literature surveys are used to examine whether general relationships between $G_{Sref}$, $h$, and $A_S/A_L$ emerge that are sufficiently strong to eclipse inter-specific variation in $\Psi_{leaf}$ and $k_s$.

Mechanistic relationships between the parameters controlling $G_S$ and remotely-sensed features of canopy architecture (such as $h$), if present, could improve biosphere-atmosphere mass and energy exchange estimates at large spatial scales. To my knowledge, no attempt has been made to determine whether such generic relationships exist between measurable features of hydraulic architecture and canopy conductance among diverse species at the level of simplicity that permits incorporation into coarse-scale models. On the other hand, relationships between canopy conductance and features of canopy architecture have been well documented within species. A predictable decrease in both leaf-level and mean canopy stomatal conductance with canopy height has been reported for a range of species, including \textit{Fagus sylvatica} (Schäfer et al., 2000), \textit{Picea abies} (Köstner et al., 2002), \textit{Pinus palustris} (Addington et al., 2006), \textit{Pinus pinaster} (Delzon et al., 2004), \textit{Pinus ponderosa} (Ryan et al.,
In many cases, this decrease is attributed to increased hydraulic resistance associated with increased path length. However, several of these studies also suggest that $A_S/A_L$ is another important determinant of $G_S$ (Schäfer et al., 2000; Becker et al., 2000; Mencuccini and Bonosi, 2001; Delzon et al., 2004), and in some cases alterations in $A_S/A_L$ can nearly compensate for height or physiologically based reductions in $G_S$ (Barnard and Ryan, 2003). It is therefore likely that the most parsimonious generic model of canopy conductance accounting for readily measurable features of hydraulic architecture must consider, at minimum, $A_S/A_L$ and $h$. This investigation was made to assess the performance of such a model over a wide range of climatic regimes and species.

3.3 Methods

Two independent literature surveys were conducted. The first survey was designed to explore inter-specific variation between $G_{Sref}$, $h$, and $A_S/A_L$. The second survey was used to determine the extent of inter-specific variability in $\alpha_S$ (and hence, $A_S/A_L$), and to evaluate whether such variations can be related to climate controls, phylogenetic similarity, or other ecosystem features.

3.3.1 Survey 1 – relationships between $G_{Sref}$, $h$, and $A_S/A_L$

Published estimates of $h$ and $G_{Sref}$ were obtained and analyzed for 42 closed-canopy forest ecosystems representing a wide range of species from boreal to tropical climates (Survey 1, Table 3.1). Estimates of $A_S/A_L$ were available for 29 of these sites. These studies relied on canopy transpiration obtained by either sap-flux or eddy covariance methodologies, averaged over a range of time scales from half-hourly to daily. Typically, leaf-scale canopy conductance was derived in these studies from estimates of transpiration per unit ground area ($T$) and $D$ using (Montheith and
Unsworth, 1990):

\[ G_S = \frac{a_V \cdot T}{D \cdot LAI}, \]

(3.7)

where \( a_V \) is a temperature-dependent constant derived from the latent heat of vaporization, the specific heat capacity of dry air, mean air density, and the psychrometric constant. In the case of the six eddy-covariance estimates, measures were taken at each site to ensure that conductance was derived from measured water vapor fluxes that did not include a significant contribution from soil evaporation. In the case of the \textit{Populus tremuloides} and \textit{Pinus radiata} canopies, soil evaporation was measured independently of whole-canopy evaporation using lysimeters. In the \textit{Picea mariana} stand, soil and sub-canopy evapotranspiration were measured with a below-canopy eddy-covariance system. In the 6.8 m \textit{Pinus taeda} stand, \( G_S \) estimated from whole canopy evapotranspiration fluxes and from sap-flux data responded similarly to \( D \), suggesting that the eddy-covariance evapotranspiration fluxes in this canopy were driven primarily by transpiration. And finally, transpiration in the 16 m \textit{Pinus taeda} stand and the mixed deciduous forest were partitioned from the measured evapotranspiration fluxes using a simple radiation transfer model as described in Stoy et al. (2006a).

For sites with high leaf area, it is well known that not all the foliage contributes to transpiration. Because total conductance rates are normalized by the measured \( LAI \) to obtain \( G_S \) rather than the \( LAI \) contributing to stand transpiration, an adjustment is necessary for sites with high \( LAI \). The \( LAI \) (and hence the reference conductance rates) were corrected for sites with exceptionally high (i.e. \( LAI \geq 8 \)) by multiplying by a factor \( f = LAI/8 \). This correction is similar to that suggested by Granier et al. (2000) though the correction was only implemented for sites with \( LAI \geq 8 \) (as opposed to \( LAI \geq 6 \)) because this is roughly the value of \( LAI \) at which the fraction of absorbed radiation in the canopy reaches 95\% during midday hours when it is
Table 3.1 Summary of studies used to assess the relationship between reference canopy conductance ($G_{S_{ref}}$, mmol m$^{-2}$ s$^{-1}$), canopy height ($h$, m), and sapwood-to-leaf area ratio ($A_{S}/A_{L}$, cm$^2$ m$^{-2}$). 'Species' denotes the dominant species at each site. $T_M$ is mean annual temperature ($^\circ$C), and $LAI$ is leaf area index (m$^2$ m$^{-2}$). ‘E’ denotes eddy-covariance measurements, and ‘S’ denotes sap-flux measurements. In the case of mixed stands, family type is assigned based on the phylogeny of the dominant species in the stand.

<table>
<thead>
<tr>
<th>Species</th>
<th>Location</th>
<th>Family</th>
<th>$T_M$</th>
<th>$h$</th>
<th>$G_{S_{ref}}$</th>
<th>$LAI$</th>
<th>$A_{S}/A_{L}$</th>
<th>method</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Boreal Forests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Picea abies</em></td>
<td>64.12 N, 19.27 E</td>
<td>Pinaceae</td>
<td>2</td>
<td>9.7</td>
<td>49.13</td>
<td>6</td>
<td>4.9</td>
<td>S</td>
<td>Ward et al. (2008)</td>
</tr>
<tr>
<td><em>Picea abies</em></td>
<td>60.08 N, 17.48 E</td>
<td>Pinaceae</td>
<td>5.5</td>
<td>23</td>
<td>180</td>
<td>4.5</td>
<td>S</td>
<td>Cienciala et al. (1997)</td>
<td></td>
</tr>
<tr>
<td><em>Populus temuloides</em></td>
<td>53.63 N, 106.20 W</td>
<td>Salicaceae</td>
<td>0.4</td>
<td>22</td>
<td>134</td>
<td>3.3</td>
<td>11.3</td>
<td>E</td>
<td>Blanken and Black (2004)</td>
</tr>
<tr>
<td><em>Picea mariana</em></td>
<td>55.88 N, 90.30 W</td>
<td>Pinaceae</td>
<td>0.8</td>
<td>9</td>
<td>55</td>
<td>7.5</td>
<td>2.5</td>
<td>S</td>
<td>Ewers et al. (2005)</td>
</tr>
<tr>
<td><em>Picea mariana</em></td>
<td>55.88 N, 90.30 W</td>
<td>Pinaceae</td>
<td>0.8</td>
<td>10</td>
<td>42</td>
<td>6.1</td>
<td>2.1</td>
<td>S</td>
<td>Ewers et al. (2005)</td>
</tr>
<tr>
<td><em>Picea mariana</em></td>
<td>55.88 N, 98.48 W</td>
<td>Pinaceae</td>
<td>-3.2</td>
<td>12</td>
<td>35</td>
<td>4.6</td>
<td>2.2</td>
<td>E</td>
<td>Bartlett et al. (2003)</td>
</tr>
<tr>
<td><em>Pinus sylvestris</em></td>
<td>60.72 N, 89.13 E</td>
<td>Pinaceae</td>
<td>5.5</td>
<td>17.4</td>
<td>82</td>
<td>5.0</td>
<td>7.1</td>
<td>S</td>
<td>Poyatos et al. (2007)</td>
</tr>
<tr>
<td><em>Pinus sylvestris</em></td>
<td>60.08 N, 17.48 E</td>
<td>Pinaceae</td>
<td>5.5</td>
<td>26.8</td>
<td>33</td>
<td>4.5</td>
<td>10.1</td>
<td>S</td>
<td>Poyatos et al. (2007)</td>
</tr>
<tr>
<td><strong>Temperate Forests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Abies bornmulleriana</em></td>
<td>48.73 N, 6.23 E</td>
<td>Pinaceae</td>
<td>9.6</td>
<td>11</td>
<td>75</td>
<td>8.9</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td><em>Crataegus monogyna</em></td>
<td>51.6 N, 1.7 W</td>
<td>Rosaceae</td>
<td>9.5</td>
<td>4</td>
<td>241</td>
<td>4.8</td>
<td>8.8</td>
<td>S</td>
<td>Herbst et al. (2007)</td>
</tr>
<tr>
<td><em>Cryptomeria japonica D.</em></td>
<td>33.13 N, 130.72 E</td>
<td>Cupressaceae</td>
<td>15</td>
<td>22</td>
<td>29</td>
<td>5.4</td>
<td>6.7</td>
<td>S</td>
<td>Kumagai et al. (2008)</td>
</tr>
<tr>
<td><em>Cryptomeria japonica D.</em></td>
<td>33.13 N, 130.72 E</td>
<td>Cupressaceae</td>
<td>15</td>
<td>32</td>
<td>39</td>
<td>5.7</td>
<td>8.1</td>
<td>S</td>
<td>Kumagai et al. (2008)</td>
</tr>
<tr>
<td><em>Fagus sylvatica</em></td>
<td>48.2 N, 7.25 E</td>
<td>Fagaceae</td>
<td>9.8</td>
<td>22.5</td>
<td>75</td>
<td>5.7</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td><em>Fagus sylvatica</em></td>
<td>48.67 N, 7.08 E</td>
<td>Fagaceae</td>
<td>9.2</td>
<td>14</td>
<td>87</td>
<td>5.7</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td><em>Fagus sylvatica</em></td>
<td>49.87 N, 10.45 W</td>
<td>Fagaceae</td>
<td>6</td>
<td>23</td>
<td>83</td>
<td>6.2</td>
<td>3.9</td>
<td>S</td>
<td>Schäfer et al. (2000)</td>
</tr>
<tr>
<td>Mixed deciduous</td>
<td>33.93 N, 79.13 W</td>
<td>Juglandaceae</td>
<td>15.5</td>
<td>23</td>
<td>67</td>
<td>5.5</td>
<td>5.4</td>
<td>S</td>
<td>Pataki and Oren (2003)</td>
</tr>
<tr>
<td>Mixed deciduous</td>
<td>46.24 N, 89.35 W</td>
<td>Aceraceae</td>
<td>3.9</td>
<td>22</td>
<td>32</td>
<td>7.5</td>
<td>2.6</td>
<td>S</td>
<td>Tang et al. (2006)</td>
</tr>
<tr>
<td>Mixed deciduous</td>
<td>33.93 N, 79.13 W</td>
<td>Juglandaceae</td>
<td>15.5</td>
<td>25</td>
<td>93</td>
<td>6.1</td>
<td>5.4</td>
<td>E</td>
<td>Stoy et al. (2006a)</td>
</tr>
<tr>
<td>Mixed deciduous</td>
<td>51.79 N, 1.3 W</td>
<td>Juglandaceae</td>
<td>9.7</td>
<td>21</td>
<td>109</td>
<td>3.6</td>
<td>0.0</td>
<td>S</td>
<td>Herbst et al. (2008)</td>
</tr>
<tr>
<td>Mixed deciduous</td>
<td>51.45 N, 1.27 W</td>
<td>Fagaceae</td>
<td>10.9</td>
<td>22</td>
<td>82</td>
<td>3.9</td>
<td>0.0</td>
<td>S</td>
<td>Herbst et al. (2008)</td>
</tr>
<tr>
<td><em>Quercus alba</em></td>
<td>35.87 N, 80.00 W</td>
<td>Fagaceae</td>
<td>15.5</td>
<td>25</td>
<td>40</td>
<td>3.1</td>
<td>1.1</td>
<td>S</td>
<td>Oren and Pataki (2001)</td>
</tr>
<tr>
<td><em>Picea abies</em></td>
<td>48.73 N, 6.23 E</td>
<td>Pinaceae</td>
<td>9.6</td>
<td>11</td>
<td>66</td>
<td>9.5</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td>Species</td>
<td>Latitude, Longitude</td>
<td>Family</td>
<td>Min GPP</td>
<td>Max GPP</td>
<td>Avg GPP</td>
<td>StdDev</td>
<td>Tn</td>
<td>Estimate</td>
<td>Source</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>--------</td>
<td>-----</td>
<td>----------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Picea abies</td>
<td>48.2 N, 7.25 E</td>
<td>Pinaceae</td>
<td>6</td>
<td>13</td>
<td>93</td>
<td>6.1</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>50.15 N, 11.87 E</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>16.1</td>
<td>66</td>
<td>5.3</td>
<td>3.8</td>
<td>S</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>50.15 N, 11.87 E</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>14.7</td>
<td>84</td>
<td>6.4</td>
<td>3.6</td>
<td>S</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>50.15 N, 11.87 E</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>17.8</td>
<td>62</td>
<td>7.1</td>
<td>3.7</td>
<td>S</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>50.15 N, 11.87 E</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>24.1</td>
<td>44</td>
<td>7.9</td>
<td>2.6</td>
<td>S</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>50.15 N, 11.87 E</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>25.7</td>
<td>56</td>
<td>7.6</td>
<td>2.4</td>
<td>S</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>50.15 N, 11.87 E</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>25.2</td>
<td>31</td>
<td>6.5</td>
<td>2.1</td>
<td>S</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Pinus pinaster</td>
<td>44.70 N, 0.77 W</td>
<td>Pinaceae</td>
<td>9.8</td>
<td>12</td>
<td>104</td>
<td>4.4</td>
<td>8.4</td>
<td>S</td>
<td>Granier and Loustau (1994)</td>
</tr>
<tr>
<td>Pinus pinaster</td>
<td>44.08 N, 0.08 W</td>
<td>Pinaceae</td>
<td>12.5</td>
<td>18</td>
<td>87</td>
<td>12.5</td>
<td>5.7</td>
<td>S</td>
<td>Porte and Loustau (1998)</td>
</tr>
<tr>
<td>Pinus taeda</td>
<td>34.80 N, 72.20 W</td>
<td>Pinaceae</td>
<td>15.5</td>
<td>6.8</td>
<td>154</td>
<td>3.5</td>
<td>6.8</td>
<td>E</td>
<td>Lai et al. (2002)</td>
</tr>
<tr>
<td>Pinus taeda</td>
<td>33.93 N, 79.13 W</td>
<td>Pinaceae</td>
<td>15.5</td>
<td>16</td>
<td>113</td>
<td>4.5</td>
<td>8.2</td>
<td>E</td>
<td>Stoy et al. (2006a)</td>
</tr>
<tr>
<td>Pinus radiata</td>
<td>42.87 S, 172.75 E</td>
<td>Pinaceae</td>
<td>10.8</td>
<td>8</td>
<td>75</td>
<td>6.5</td>
<td>0.0</td>
<td>E</td>
<td>Arneth et al. (1999)</td>
</tr>
<tr>
<td>Populus trichocarpa</td>
<td>46.17 N, 118.47 W</td>
<td>Salicaceae</td>
<td>12.3</td>
<td>8</td>
<td>148</td>
<td>9.5</td>
<td>3.3</td>
<td>S</td>
<td>Kim et al. (2008)</td>
</tr>
<tr>
<td>Quercus petraea</td>
<td>48.7 N, 6.4 E</td>
<td>Fagaceae</td>
<td>9.6</td>
<td>15</td>
<td>95</td>
<td>6.0</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
</tbody>
</table>

**Tropical Forests**

<table>
<thead>
<tr>
<th>Species</th>
<th>Latitude, Longitude</th>
<th>Family</th>
<th>Min GPP</th>
<th>Max GPP</th>
<th>Avg GPP</th>
<th>StdDev</th>
<th>Tn</th>
<th>Estimate</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eperua falcata</td>
<td>5.2 N, 52.7 W</td>
<td>Fabaceae</td>
<td>25.8</td>
<td>10</td>
<td>43</td>
<td>10.8</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td>Eucalyptus saligna</td>
<td>19.84 N, 155.12 W</td>
<td>Myrtaceae</td>
<td>21</td>
<td>7</td>
<td>40</td>
<td>4.9</td>
<td>0.7</td>
<td>S</td>
<td>Barnard and Ryan (2003)</td>
</tr>
<tr>
<td>Eucalyptus saligna</td>
<td>19.84 N, 155.12 W</td>
<td>Myrtaceae</td>
<td>21</td>
<td>26</td>
<td>37</td>
<td>5.1</td>
<td>1.8</td>
<td>S</td>
<td>Barnard and Ryan (2003)</td>
</tr>
<tr>
<td>Goupia glabra</td>
<td>5.2 N, 52.7 W</td>
<td>Goupiaceae</td>
<td>25.8</td>
<td>15</td>
<td>74</td>
<td>4.3</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td>Mixed tropical</td>
<td>5.2 N, 52.7 W</td>
<td>Fabaceae</td>
<td>25.9</td>
<td>33</td>
<td>57</td>
<td>8.6</td>
<td>1.5</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
<tr>
<td>Simarouba amara</td>
<td>5.2 N, 52.7 W</td>
<td>Simaroubaceae</td>
<td>25.8</td>
<td>4.7</td>
<td>108</td>
<td>3.5</td>
<td>0.0</td>
<td>S</td>
<td>Granier et al. (2000)</td>
</tr>
</tbody>
</table>

Table 3.1:
modeled from Beer’s Law (Campbell and Norman, 1998).

The reported values of $G_{Sref}$ obtained from the literature were estimated using a range of analytical procedures, including boundary line analyses, optimization routines, and data binning. In all cases, the extracted value represents the authors’ estimate of the conductance rate at the reference $D$ of 1 kPa under conditions of non-limiting light and soil moisture content. In this analysis, $G_{Sref}$ is expressed in mmol m$^{-2}$ s$^{-1}$. Reference conductance measurements presented in units of mm s$^{-1}$ in the original source were converted using the molar density of water vapor in air at 25 °C.

The analysis is restricted to closed canopies because trees in open canopies are more likely to have a conical or complicated branch architecture, which weakens the link between $h$ and mean path length. Data from manipulation experiments was also excluded because sapwood permeability and $A_S/A_L$ may respond to abrupt changes in nutrient or light regimes, achieved through fertilization (Ewers et al., 2005; Phillips et al., 2001), stand density reduction (Simonin et al., 2006), CO$_2$ enrichment (Atwell et al., 2003; Pataki et al., 2000), and foliage removal (Pataki et al., 1998), and the adjustment to new conditions may take several years. In nearly all of these studies, $G_{Sref}$ is normalized by maximum LAI in the growing season, and $A_S/A_L$ represents the ratio of sapwood-to-leaf area at breast height to projected leaf area during the growing season. However, I did not exclude studies that reported estimates of these parameters derived from total leaf area as opposed to LAI, or studies for which sapwood area estimates are taken from a different height (Herbst et al., 2007), to maximize the sample size in Table 3.1. No other exclusionary criteria were employed in this survey.

The variables of interest were treated as canopy averages in these surveys. In the cases where data was reported for individual trees or species, canopy averages were calculated by weighting individual- or species-specific values according to their LAI.
3.3.2 *Allometric equations for* $A_S/A_L$

In a second literature survey, the slope and intercept of the change in $A_S/A_L$ with $h$ were compiled from studies on 21 closed canopy forest ecosystems (Survey 2, Table 3.2), representing different species growing in a broad range of climates. Estimates of canopy-averaged values of $A_S/A_L$ and $h$ along chronosequence stages, as well as whole-tree estimates of $A_S/A_L$ for trees of different heights in the same stand, were considered. The same exclusionary criteria employed for Survey 1 were employed for Survey 2. Survey 2 is similar to a survey conducted by McDowell et al. (2002a) yet less than a quarter of the studies cited in Table 3.2 are common to both surveys. However, in this study, I expanded considerably the sample size and the number of sites which have a negative relationship between $A_S/A_L$ and $h$ (i.e. negative $a_S$).

3.3.3 *Statistical tests and optimization*

Statistical performance indicators such as the correlation coefficient ($r^2$) and $t$-statistics for slope significance (i.e. $P$) were performed in MATLAB version 6.0. Because correlation coefficients are often compared between datasets of different sample size in this study, adjusted $r^2$ is used. Unless otherwise stated, slope significance was interpreted using two-tailed $t$-tests with a null hypothesis of zero slope. When necessary, nonlinear optimization was performed in MATLAB.

3.4 *Results*

3.4.1 *Changes in* $G_{Sref}$ *with* $A_S/A_L$ *and* $h$

Using Equation 3.1 along with simplifications leading to Equation 3.4, $G_{Sref}$ was shown to be analytically related to the product of $A_S/A_L$ and $h^{-1}$, a finding that appears to be accurate across the 29 sites for which all three variables were available (Survey 1, Table 3.1, Figure 3.2). A simple linear regression of these variables gives:
<table>
<thead>
<tr>
<th>Species</th>
<th>Family</th>
<th>$T_M$</th>
<th>$P_M$</th>
<th>min $h$</th>
<th>min $A_S/A_L$</th>
<th>$\alpha_S$</th>
<th>$\beta_S$</th>
<th>$n$</th>
<th>Type</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abies balsamea</td>
<td>Pinaceae</td>
<td>4</td>
<td>1000</td>
<td>2</td>
<td>2.61</td>
<td>-0.14</td>
<td>3.61</td>
<td>56</td>
<td>C</td>
<td>Coyea and Margolis (1992)</td>
</tr>
<tr>
<td>Abies balsamea</td>
<td>Pinaceae</td>
<td>6.6</td>
<td>1060</td>
<td>7.6</td>
<td>0.76</td>
<td>0.13</td>
<td>0.39</td>
<td>3</td>
<td>T</td>
<td>Gilmore et al. (1996)</td>
</tr>
<tr>
<td>Abies lasiocarpa</td>
<td>Pinaceae</td>
<td>2</td>
<td>720</td>
<td>3.8</td>
<td>0.313</td>
<td>0.06</td>
<td>0.24</td>
<td>9</td>
<td>T</td>
<td>Sala (2006)</td>
</tr>
<tr>
<td>Eucalyptus delegatensis</td>
<td>Myrtaceae</td>
<td>9.5</td>
<td>1400</td>
<td>3</td>
<td>2.62</td>
<td>-0.04</td>
<td>3.25</td>
<td>23</td>
<td>T</td>
<td>Mokany et al. (2003)</td>
</tr>
<tr>
<td>Eucalyptus saligna</td>
<td>Myrtaceae</td>
<td>21</td>
<td>4000</td>
<td>7</td>
<td>0.68</td>
<td>0.06</td>
<td>0.27</td>
<td>2</td>
<td>C</td>
<td>Barnard and Ryan (2003)</td>
</tr>
<tr>
<td>Fagus sylvatica</td>
<td>Fagaceae</td>
<td>7.5</td>
<td>750</td>
<td>11</td>
<td>2.83</td>
<td>0.16</td>
<td>1.5</td>
<td>9</td>
<td>T</td>
<td>Schäfer et al. (2000)</td>
</tr>
<tr>
<td>Larix occidentalis</td>
<td>Pinaceae</td>
<td>7.2</td>
<td>430</td>
<td>11</td>
<td>2.44</td>
<td>-0.05</td>
<td>4.2</td>
<td>11</td>
<td>T</td>
<td>Sala (2006)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>Pinaceae</td>
<td>5.8</td>
<td>1100</td>
<td>14.7</td>
<td>3.55</td>
<td>-0.13</td>
<td>5.8</td>
<td>6</td>
<td>C</td>
<td>Alsheimer et al. (1998)</td>
</tr>
<tr>
<td>Picea abies</td>
<td>Pinaceae</td>
<td>2</td>
<td>600</td>
<td>8.73</td>
<td>3.27</td>
<td>-0.72</td>
<td>11.47</td>
<td>6</td>
<td>T</td>
<td>Ward et al. (2008)</td>
</tr>
<tr>
<td>Picea mariana</td>
<td>Pinaceae</td>
<td>9.8</td>
<td>440</td>
<td>2.8</td>
<td>0.8</td>
<td>-0.41</td>
<td>6.1</td>
<td>19</td>
<td>T</td>
<td>Ewers et al. (2005)</td>
</tr>
<tr>
<td>Picea sitchensis</td>
<td>Pinaceae</td>
<td>9.3</td>
<td>850</td>
<td>4.4</td>
<td>1.82</td>
<td>0.03</td>
<td>3.6</td>
<td>6</td>
<td>C</td>
<td>Tobin et al. (2006)</td>
</tr>
<tr>
<td>Pinus albicaulis</td>
<td>Pinaceae</td>
<td>2</td>
<td>720</td>
<td>3.5</td>
<td>2.28</td>
<td>-0.28</td>
<td>7.63</td>
<td>14</td>
<td>T</td>
<td>Monserud and Marshall (1999)</td>
</tr>
<tr>
<td>Pinus monticola</td>
<td>Pinaceae</td>
<td>6.6</td>
<td>810</td>
<td>5</td>
<td>2.98</td>
<td>0.08</td>
<td>2.89</td>
<td>21</td>
<td>T</td>
<td>Monserud and Marshall (2001)</td>
</tr>
<tr>
<td>Pinus ponderosa</td>
<td>Pinaceae</td>
<td>6.6</td>
<td>810</td>
<td>4</td>
<td>5.71</td>
<td>0.16</td>
<td>9.33</td>
<td>22</td>
<td>T</td>
<td>Monserud and Marshall (2001)</td>
</tr>
<tr>
<td>Pinus sitchensis</td>
<td>Pinaceae</td>
<td>7.2</td>
<td>430</td>
<td>13.2</td>
<td>5.97</td>
<td>0.01</td>
<td>5.74</td>
<td>11</td>
<td>T</td>
<td>Sala (2006)</td>
</tr>
<tr>
<td>Pinus sylvestris</td>
<td>Pinaceae</td>
<td>10</td>
<td>550</td>
<td>8</td>
<td>9.31</td>
<td>0.17</td>
<td>7.6</td>
<td>5</td>
<td>C</td>
<td>Magnani et al. (2000)</td>
</tr>
<tr>
<td>Pinus sylvestris</td>
<td>Pinaceae</td>
<td>6.5</td>
<td>1215</td>
<td>4</td>
<td>5.77</td>
<td>0.22</td>
<td>7.86</td>
<td>19</td>
<td>T</td>
<td>Martinez-Vilalta et al. (2007)</td>
</tr>
<tr>
<td>Pseudotsuga menziesii</td>
<td>Pinaceae</td>
<td>8.7</td>
<td>2500</td>
<td>15</td>
<td>1.93</td>
<td>0.01</td>
<td>1.7</td>
<td>3</td>
<td>C</td>
<td>Addington et al. (2006)</td>
</tr>
<tr>
<td>Pseudotsuga menziesii</td>
<td>Pinaceae</td>
<td>6.6</td>
<td>810</td>
<td>6</td>
<td>2.94</td>
<td>0.01</td>
<td>3.02</td>
<td>23</td>
<td>T</td>
<td>Lefsky et al. (1999)</td>
</tr>
<tr>
<td>Pseudotsuga menziesii</td>
<td>Pinaceae</td>
<td>7.2</td>
<td>434</td>
<td>11</td>
<td>1.95</td>
<td>-0.17</td>
<td>2.04</td>
<td>17</td>
<td>T</td>
<td>Phillips et al. (1996)</td>
</tr>
<tr>
<td>Quercus garryana</td>
<td>Fagaceae</td>
<td>11</td>
<td>1100</td>
<td>10</td>
<td>4.34</td>
<td>-0.11</td>
<td>5.4</td>
<td>2</td>
<td>C</td>
<td>Phillips et al. (2003a)</td>
</tr>
</tbody>
</table>

Table 3.2: Summary of studies in closed-canopy forests used to assess the relationship between sapwood-to-leaf area ratio ($A_S/A_L$) and mean canopy height ($h$). $T_M$ and $P_M$ are mean annual temperature and precipitation, respectively. Min $h$ (m) and min $A_S/A_L$ ($\text{cm}^2 \text{ m}^{-2}$) are the values associated with the shortest tree in each dataset. $\alpha_S$ ($\text{cm}^2 \text{ m}^{-3}$) and $\beta_S$ ($\text{cm}^2 \text{ m}^{-3}$) are the slope and intercept of the linear relationship between $A_S/A_L$ and $h$ (see Equation 3.5). The number of individual measurements used to derive the relationships is denoted by $n$. The data types are: 1) C: whole-canopy measurements, and 2) T: individual tree measurements.
\[ G_{\text{Sref}} = 98.2A_S/A_Lh + 37.3, \]  

(3.8)

with \( r^2 = 0.75 \) and \( P < 0.0001 \). Separating the relative importance of \( A_S/A_L \) and \( h^{-1} \), approximately 27% of the variability in \( G_{\text{Sref}} \) is driven by \( A_S/A_L \) \( (P < 0.01) \) and 46% is driven by \( h^{-1} \) \( (P < 0.0001) \). The relationship is also quite strong when reference canopy rates uncorrected for high \( \text{LAI} \) are considered (inset to Figure 3.2, \( r^2 = 0.73, P < 0.0001 \)).

The sites in the above analysis included 19 temperate, seven boreal and three tropical forest ecosystems. The small sample size of boreal and tropical forest sites prevents this relationship from being analyzed within each of these climatically distinct subsets. However, in temperate sites, the slope of the relationship \( (G_{\text{Sref}} = 95.8A_S/A_L/h + 43.2, r^2=0.92) \) is not statistically distinguishable from the slope derived with data from all three climate zones \( (P = 0.81) \).

Among the 29 sites, seven are dominated by \textit{Picea abies}, three are dominated by \textit{Picea mariana}, and two each are dominated by \textit{Cryptomeria Japonica}, \textit{Pinus pinaster}, \textit{Pinus taeda}, \textit{Eucalyptus saligna}, and \textit{Pinus sylvestris}. To assess the influence of replicates of single species, a replication analysis procedure proposed by McDowell et al. (2002a) was adopted. Specifically, the analysis was repeated for 672 unique combinations of sites such that no more than one site dominated by each species was included. Each combination resulted in a positive slope (ranging 92.5 to 98.4 mmol m\(^{-1}\) s\(^{-1}\)) that remained statistically different from zero \( (P < .0001 \text{ for all combinations}) \). Furthermore, none of the slopes differed significantly from the slope derived from the entire dataset \( (P > 0.6 \text{ for all combinations}) \).

A weak relationship between \( G_{\text{Sref}} \) and \( h^{-1} \) emerged when analyzing all 42 datasets presented in Table 3.1 \( (r^2=0.24, P<0.001) \). \( G_{\text{Sref}} \) and \( h^{-1} \) were more significantly correlated when temperate sites were analyzed separately. Across temperate
Figure 3.2: The relationship between reference conductance ($G_{Sref}$) and the product of the ratio of sapwood-to-leaf area ($A_S/A_L$) and the inverse of canopy height ($h^{-1}$). The solid line is determined from least squares regression using all data, and the dotted line is the least squares regression for temperate forests only. Open symbols denote canopies dominated by species known to have a decreasing relationship between $A_S/A_L$ and $h$. The insert shows the same relationship for estimates of $G_{Sref}$ uncorrected for high LAI as described in Section 3.3.

sites, reference conductance increased strongly with $h^{-1}$ ($r^2=0.68$, $p<0.0001$, Figure 3.3a). Again adopting the analysis replication procedure (giving 192 unique combinations), the relationship between $G_{Sref}$ and $h^{-1}$ was significant for all combinations of sites in which only one stand of each species was represented ($P < 0.001$ for all combinations). This relationship, however, is driven strongly by the data from the 4-m hedgerow stand (Table 3.1). Excluding this site from the analysis, the increase
Figure 3.3: Reference canopy conductance ($G_{Sref}$) vs. the inverse of canopy height ($h^{-1}$) for (a) temperate, (b) tropical, and (c) boreal species. Symbols are the same as Figure 2. Regression lines are not shown for tropical and boreal sites as no significant relationships between $G_{Sref}$ and $h^{-1}$ emerged for these small samples.

In $G_{Sref}$ with $h^{-1}$ was significantly different from zero (at the 95% confidence level) for all combinations that included the 6.8 m *Pinus taeda* stand.

Among tropical species, $h^{-1}$ explained 19% of the variance in $G_{Sref}$, although the slope is not statistically significant (Figure 3.3b, $P = 0.17$). The tropical subset includes two *Eucalyptus saligna* stands, but repeating the analysis using one or the other of these sites resulted in a derived slope that was statistically indistinguishable from the slope calculated from all tropical sites.

$G_{Sref}$ decreased weakly and insignificantly with $h^{-1}$ among the boreal sites (Figure 3.3c, $r^2 = 0.18$, $P = 0.17$) though the decrease is significant for some combinations.
of boreal sites that included only one representation of each species. This negative relationship is driven by reference conductance rates of *Picea mariana* (i.e. - the three boreal sites with the highest value of $A_S/A_L/h$). *Picea mariana* has a strongly decreasing $\alpha_S$ (Ewers et al., 2005), and would be expected to have relatively low reference conductance rates.

Roughly 50% of the studies considered in Survey 1 are from the *Pinaceae* family. Therefore, for the significant relationships that emerged from this analysis (i.e. Figure 3.2, Figure 3.3a), I conducted two additional tests to assess the impact of phylogenetic similarities among the ecosystems: 1) an additional replication analysis procedure was adopted whereby the relationships were assessed for unique combinations of sites such that no more than one species from each family was represented, and 2) the relationships were derived independently for angiosperms and gymnosperms. For the relationship between $G_{Sref}$ and $A_S/A_L/h$ shown in Figure 3.2, all 512 unique combinations resulted in a statistically significant slope ($P < 0.001$) with a high degree of correlation ($r^2 = 0.79$ to 0.91). The correlation for the relationship derived with angiosperms alone ($G_{Sref} = 45.0A_S/A_L/h + 71.1$) improved significantly when compared to the relationship derived with gymnosperms alone ($r^2=0.92$ and 0.78, respectively), though this higher correlation is driven strongly by the reference canopy rate in the 4-m hedgerow (an angiosperm site). For the relationship between $G_{Sref}$ and $1/h$ among temperate forests (Figure 3.3a), all 1008 unique combinations resulted in statistically significant slopes ($P < 0.01$). The amount of variance in $G_{Sref}$ explained by $1/h$ is higher for angiosperms alone ($r^2 = 0.92$), though again, this relationship is driven strongly by the hedgerow.

Finally, because reference conductance rates have previously been shown to vary with leaf area within species, I also assessed the generality of this relationship. Total reference conductance (i.e. reference conductance per unit ground area) should increase with $LAI$; however, due to the saturation of canopy light absorption at high
Figure 3.4: Reference canopy conductance ($G_{\text{Sref}}$) as a function of leaf area index ($LAI$) for all sites in Table 3.1. Symbols are the same as Figure 3.2.

$LAI$, reference conductance per unit leaf area should decrease with $LAI$ in dense stands. A significant but very weak linear negative relationship between $G_{\text{Sref}}$ and $LAI$ was observed based on the 42 sites of Survey 1 ($r^2=0.08$, $P < 0.05$, Figure 3.4), with correlation improving slightly for the relationship between $G_{\text{Sref}}$ and $\ln(LAI)$ ($r^2 = 0.10$).

3.4.2 Relationship between $A_S/A_L$ and $h$

The linear relationship between $A_S/A_L$ and $h$ compiled from the literature varied considerably among the 21 sites considered in Survey 2 (Table 3.2). A majority of the studies reported a positive linear relationship, though due to the pres-
ence of some strongly negative slopes, the overall mean values were \( \langle \alpha_S \rangle = -0.03 \) and \( \langle \beta_S \rangle = 4.3 \), with standard deviations of \( \sigma(\alpha_S) = 0.18 \) and \( \sigma(\beta_S) = 2.65 \), respectively. To determine whether this variation is sufficient to explain the variation observed in the general relationship between \( G_{Sref} \) and \( h \) (Figure 3.3), the quantity \( G_{Sref} \approx (\langle \alpha_S \rangle h + \langle \beta_S \rangle) / h \) was referenced to the conductance data by minimizing the standard error between this quantity and the measurements (Figure 3.5). This model clearly accounts for very little of the variability; however, more than 70% of the data points fall within the range of expectation bounded by \( G_{Sref} \approx (\langle \alpha_S \rangle \pm \sigma(\alpha_S)) h + \langle \beta_S \rangle) / h \) (shaded area in Figure 3.5), suggesting that much of the observed variability in \( G_{Sref} \) may be explained by the large variations of \( \alpha_S \) among species.

The mean values \( \langle \alpha_S \rangle \) and \( \langle \beta_S \rangle \) did not change significantly when the analysis was repeated to eliminate multiple data sets of one species. Furthermore, the mean values of \( \alpha_S \) and \( \beta_S \) for relationships derived using whole-canopy values of \( A_S/A_L \) and \( h \) among chronosequences (\( \langle \alpha_{S,chr} \rangle = -0.015 \), \( \langle \beta_{S,chr} \rangle = 4.0 \)) were statistically indistinguishable from the mean values of \( \alpha_S \) and \( \beta_S \) for relationships derived using measurements of \( A_S/A_L \) and \( h \) on individual trees within a single stand \( (\langle \alpha_{S,stand} \rangle = -0.030 \), \( \langle \beta_{S,stand} \rangle = 4.3 \) using a t-test for differences between the means assuming unknown but equal variances.

The slope factor \( \alpha_S \) was not related to mean annual precipitation (which can be considered a proxy for soil water availability) across sites, consistent with previous inter-specific observations (McDowell et al., 2002a) and with the previous finding that \( \alpha_S \) was indistinguishable between xeric and mesic Pinus palustris stands (Addington et al., 2006). However, \( \alpha_S \) increases significantly with the natural log of mean annual temperature \( (T_M, \text{Figure 3.6, } r^2 = 0.39, P < 0.01) \), consistent with previous observations of a significant relationship between \( T_M \) and \( \alpha_S \) among mature Pinus sylvestris stands (Mencuccini and Bonosi, 2001), though much of the variation in \( \alpha_S \)
Figure 3.5: Reference canopy conductance ($G_{S_{ref}}$) vs. canopy height ($h$) for all sites from Table 1. The dotted line represents the quantity $G_{S_{ref}} \approx (\langle \alpha_s \rangle h + \langle \beta \rangle)/h$ referenced to the conductance data by minimizing the standard error ($r^2=0.24$, $P<0.001$), where $\langle \alpha_s \rangle$ and $\langle \beta \rangle$ are the average slope and intercept, respectively of the relationships presented in Table 3.2. The shaded area represents the range of expectation bounded by $G_{S_{ref}} \approx [(\langle \alpha_s \rangle \pm \sigma(\alpha_s)) h + \langle \beta \rangle]/h$, where $\sigma(\alpha_s)$ is the standard deviation of the slopes presented in Table 3.2. Symbols are as in Figure 3.2.

is not explained by temperature.

3.5 Discussion

3.5.1 The hydraulic controls on stomatal conductance across species

In 1997, Ryan & Yoder proposed that the nearly universal declines in tree growth with forest age may be related to decreasing stomatal conductance as trees grow taller and hydraulic resistance to water flow increases with the transport path length. Since
then, numerous analysis and experiments have been conducted to test this so-called “hydraulic limitation hypothesis”. Some experiments support the hypothesis (Ryan et al., 2000; Schäfer et al., 2000; Lai et al., 2002; Köstner et al., 2002; Phillips et al., 2003b; Delzon et al., 2004; Addington et al., 2006), while others suggest that $A_S/A_L$ is more important than $h$ in controlling stomatal conductance (Becker et al., 2000; Mencuccini and Bonosi, 2001; Phillips et al., 2001; Ewers et al., 2005), and some point to the importance of age or size-related changes in physiology (Thomas and Winner, 2002; Barnard and Ryan, 2003). These results confirmed the importance of
homeostatic changes in both $h$ and $A_S/A_L$ to the whole-plant water balance. I found only a weak general relationship between reference conductance and height alone among 42 forested ecosystems representing a large number of species from a wide range of climates, although a strong relationship exists within the better represented temperate climate subset (Figure 3.3a). Adding $A_S/A_L$ to $h$ explains 75% of the variation in $G_{S_{ref}}$ among 29 sites representing a wide range of biomes (Figure 3.2). This degree of explanatory power exceeded that predicted by the theoretical arguments of Section 3.2, which projected equal influence of $k_s(\Psi_{leaf} - \rho_w gh)$ and $A_S/A_L/h$ on $G_{S_{ref}}$. That $A_S/A_L/h$ eclipses $k_s(\Psi_{leaf} - \rho_w gh)$ in terms of impact on reference conductance rates across species suggests compensatory interactions between $k_s$ and $\Psi_{leaf}$ limiting the range of $k_s(\Psi_{leaf} - \rho_w gh)$ that may exist across species, or that these interaction are mediated by height or $A_S/A_L$.

Many of the species considered in Survey 1 are phylogenetically similar, and over half are from the family Pinaceae. The significant relationships that emerged from these surveys remain relatively unchanged when only one representative of each species or family is considered in the analysis, and the dataset is more largely limited by a paucity of data from short forests as the correlations for the relationships in Figures 3.2 and 3.3a are driven strongly by the two shortest canopies (i.e. the 4 m hedgerow stand and the 6.8 m $P. \text{taeda}$ stand). Short stands, in addition to being underrepresented in this dataset, are also more subject to biases associated with equating path length to $h$. As demonstrated in Figure 3.1, neglecting rooting length in short canopies results in an overestimation of the product $A_S/A_L/h$ on the order of 10 - 20%. Conversely, canopy architecture patterns may be significantly different in shorter stands (i.e. more branching) such that $h$ may either over or under-estimate path length. While the results shown in Figure 3.2 and Figure 3.3a are robust and remain highly significant when the assumed height of these two shortest stands is altered by $\pm 2$ m, an overestimation of canopy height in these stands may suggest
a relationship between $G_{S\text{ref}}$ and $A_S/A_L/h$ or $1/h$ that is linear when a saturating function is actually a better model.

It is worth noting that the estimates of $G_{S\text{ref}}$ extracted from the literature for Survey 1 are subjective estimates determined using a range of regression and modeling procedures that vary from study to study. However, the high correlation between these estimates and $A_S/A_L/h$ suggests that the error associated with difference in methodology between the studies is relatively small.

3.5.2 Mechanisms and limits to hydraulic compensation within species

To assess the predictive ability of this model within a species, the four sites for which changes in $G_{S\text{ref}}$ and $A_S/A_L$ were reported for trees or stands of different heights were further explored. These were *Eucalyptus saligna* (Barnard and Ryan, 2003), *Fagus sylvatica* (Schäfer et al., 2000), *Picea abies* (Alsheimer et al., 1998), and *Picea mariana* (Ewers et al., 2005). Following the sensitivity analysis presented in Appendix A, the relative contribution of $\partial h/h$ and $\partial A_S/A_L/A_S/A_L$ to $\partial G_{S\text{ref}}/G_{S\text{ref}}$ is directly assessed within these species (Table 3.3). For the four datasets, the relative change in $A_S/A_L$ is insufficient to compensate for the observed reductions in conductance with increasing height. For *Eucalyptus saligna* and *Fagus sylvatica*, the ratio of the relative change in $A_S/A_L$ to the relative change in $h$ is 0.64 and 0.41, respectively. For *Picea mariana* and *Picea abies*, the observed decreases in $A_S/A_L$ with height compounds the relative decreases in $G_{S\text{ref}}$ observed in taller stands.

Spruce and fir species often exhibit negative relationships between $A_S/A_L$ and $h$ (Köstner et al., 2002; McDowell et al., 2002a; Ewers et al., 2005), which confers no known hydraulic advantage. It was proposed that this negative relationship may reflect a longer period of juvenile wood development, which has lower conductivity than latewood (Phillips et al., 1996), or increased leaf life span, which would increase nutrient recycling in poor quality sites (McDowell et al., 2002a). The latter hypoth-
esis is supported in part by the observation that $\alpha_s$ is related across species to the site quality (McDowell et al., 2002a), which reflects, among other factors, the effect of site nutrient availability on growth.

The relative rates of change shown in Table 3.3 can also be used to assess the assumptions of the proposed model for $G_{S_{ref}}$. For *Fagus sylvatica* and *Picea mariana*, the ratio of the relative change in the quantity $\Delta A_S/A_L / A_S/A_L - \Delta h/h$ is close to 1 (0.93 and 1.13, respectively), which suggests that the assumptions in this model are correct. However, the predicted change in conductance for *Picea abies* (-1.11) and *Eucalyptus saligna* (-0.93) is inconsistent with the observed relative decrease (-0.63 and -0.074, respectively), which indicates that, in some species, compensatory mechanisms other than $A_S/A_L$ and $h$ may represent important controls on reference stomatal conductance. Other compensatory changes may include height-related increases in sapwood permeability (Pothier et al., 1989), decreases in leaf water potential (McDowell et al., 2002b; Phillips et al., 2002; Barnard and Ryan, 2003), increased reliance on stored water (Phillips et al., 2003b), increased allocation to fine roots (Magnani et al., 2000), and changes in crown architecture such as increased branching and decreased stem diameter (Rust and Roloff, 2002). Data on these homeostatic mechanisms is scarce and does not support an analysis of a general relationship.

<table>
<thead>
<tr>
<th>Ecosystem</th>
<th>$\Delta G_{S_{ref}}$</th>
<th>$\Delta h$</th>
<th>$\Delta A_S/A_L$</th>
<th>$A_S/A_L - \Delta h/h$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eucalyptus saligna</em></td>
<td>-0.1</td>
<td>2.6</td>
<td>1.7</td>
<td>-0.9</td>
</tr>
<tr>
<td><em>Fagus sylvatica</em></td>
<td>-1.6</td>
<td>2.5</td>
<td>1.0</td>
<td>-1.4</td>
</tr>
<tr>
<td><em>Picea abies</em></td>
<td>-0.6</td>
<td>0.7</td>
<td>-0.4</td>
<td>-1.1</td>
</tr>
<tr>
<td><em>Picea mariana</em></td>
<td>-0.2</td>
<td>0.1</td>
<td>-0.2</td>
<td>-0.3</td>
</tr>
</tbody>
</table>

Table 3.3: The relative change in conductance ($G_S$), height ($h$), and sapwood-to-leaf area ratio ($A_S/A_L$) for the four ecosystems in Table 3.1 for which all three variables were available at various heights.
3.5.3 Variation in the rate of change of $A_S/A_L$

The primary result from Survey 1 is Equation 3.8, which shows that when $A_S/A_L$ and $h$ are measured or independently estimated, $G_{Sref}$ can be well reproduced, though $h$ alone appears to be a good predictor for temperate species. However, $A_S/A_L$ and $h$ are typically not independent within species, and may not be independent among species. Hence, Survey 2 was conducted to assess whether variations in $h$ may provide prognostic information about variations in $A_S/A_L$.

The change in sapwood-to-leaf area ratio with height varies considerably among the species of Survey 2, with the rate of change ranging from $-0.72 \text{ cm}^2 \text{ m}^{-3}$ in *Picea abies* to $0.21 \text{ cm}^2 \text{ m}^{-3}$ in *Pinus sylvestris* (Magnani et al., 2000). A mechanistic model for this variation would greatly enhance the generality of the derived relationship between $G_{Sref}$, $h$ and $A_S/A_L$, (Figure 3.2, Equation 3.8). While a significant relationship emerged from Survey 2 between $\alpha_S$ and mean annual temperature, I do not believe that this relationship is strong enough for general application at this time. In this section, some additional likely controls on height related changes in $A_S/A_L$ are discussed.

McDowell et al. (2002a) observed that in species exhibiting a positive relationship between $A_S/A_L$ and $h$, $\alpha_S$ was approximately an order of magnitude higher in vessel bearing species when compared to tracheid bearing species. In species having such positive relationships among those assembled for this analysis, the mean rate of change was only marginally higher in vessel bearing species ($\langle \alpha_{Svessel} \rangle = 0.018 \text{ cm}^2 \text{ m}^{-3}$) than tracheid bearing species ($\langle \alpha_{Stracheid} \rangle = -0.045 \text{ cm}^2 \text{ m}^{-3}$). Positive and negative values of $\alpha_S$ were reported for both tracheid and vessel bearing species, and the average rate of change for each functional type was statistically indistinguishable from the average rate of change for all species according to a t-test for differences between the means assuming unknown but equal variances (null hypoth-
esis of equivalent means). This rate of change also varies across sites occupied by the same species. For example, values of \( \alpha_S = 0.01 \) and \( \alpha_S = -0.17 \text{ m}^2 \text{ m}^{-3} \) were reported for Pseudotsuga menziesii stands, and considerable variation in \( \alpha_S \) among Pinus sylvestris and Pinus ponderosa has also been observed (see (McDowell et al., 2002a)). Thus, a simple categorization into plant functional type, or even analysis limited to a species, does not introduce much 'prognostic' utility for specifying the rate of change of \( A_S/A_L \) with height.

The lack of similarity in the sensitivity of \( A_S/A_L \) to \( h \) within plant functional types or within a species suggests that climatic controls may influence inter-site differences in \( \alpha_S \). Additionally, the absence of a strong relationship between \( G_{Sref} \) and \( h \) among all sites in the data set, but the existence of significant relationships within the temperate zone suggest that \( A_S/A_L \) reflects the prevailing climate conditions. Examination of Equation 3.3 shows that acclimation for the purpose of sustaining \( G_{Sref} \) in dry climates could be achieved through a proportional increase of \( A_S/A_L \) with \( D \). While long term average \( D \) was not available for most of the sites considered in this study, the observed relationship between \( \alpha_S \) and \( T_M \) could imply a relationship between \( \alpha_S \) and \( D \), as long-term average vapor pressure deficit and temperature are correlated across ecosystems that are not persistently water limited. In other studies, this theoretical prediction has been confirmed for Pinus sylvestris (Mencuccini and Bonosi, 2001) and other species of the genus Pinus (DeLucia et al., 2000), though no relationship between \( D \) and \( A_S/A_L \) was observed among other conifer species (i.e., Abies and Picea spp., Pseudotsuga menziesii).

Lastly, the light environment may influence the rate in which sapwood-to-leaf area ratio changes with height even within closed canopies (Oren et al., 1986). No significant differences in \( \alpha_S \) were observed between canopy-level values obtained along chronosequences of closed-canopy stands and tree-level values obtained from measurements in single stands. Because the average light environment is similar among
closed-canopy stands in a chronosequence but the light environment of individual
crowns varies considerably depending on position in the canopy, the similarity of
\( \langle \alpha_S \rangle \) in these two situations implies that the rate of change of \( A_S/A_L \) with \( h \) is not
strongly related to light availability. Indeed, values of \( \alpha_S \) for open stands (i.e. \( LAI < 3.0 \text{ m}^2 \text{ m}^{-2} \)) of \textit{Pinus ponderosa} (\( \alpha_S = 0.17 \text{ cm}^2 \text{ m}^{-3} \), Ryan et al. 2000), \textit{Pinus sylvestris} (\( \alpha_S = 0.16 \text{ cm}^2 \text{ m}^{-3} \), Delzon et al. 2004), and \textit{Pinus palustris} (\( \alpha_S = 0.21 \text{ cm}^2 \text{ m}^{-3} \), Addington et al. 2006) are well within the range of variation observed for
closed stands. In summary, future research on the sensitivity of \( A_S/A_L \) to \( h \) should
focus on the potential impacts of climate conditions and perhaps also soil nutrient
regimes, which were not explicitly considered here.

3.5.4 Broader implications for ecosystem-to-regional scale carbon and water cycle
modeling

The response of canopy conductance to rapid changes in environmental drivers is
often described with Jarvis-type multiplicative functions applied to a species-specific
reference state (here \( G_{Sref} \)). Because the Jarvis (1976) model and its variants are
widely used, much effort has been invested in deriving generic representations of the
model’s reduction functions. For example, Oren et al. (1999) showed that across
a wide range of boreal to tropical species the sensitivity of \( G_S \) to \( D \) can be well
described by the function \( f_2(D) = 1 - 0.6 \ln(D) \). Generic relationships for the light
and soil water response functions have also been developed using datasets for a broad
range of species (Granier et al., 2000). Therefore, a representation for \( G_{Sref} \) that
explains inter-site variability can be used in coordination with these generic reduction
functions to specify canopy conductance rates \textit{a priori} for a wide range of ecosystems
at a high temporal resolution.

These results suggest that differences among species in leaf physiology and the
anatomy of the transport tissue, and differences in soil properties among sites, may
exert a smaller effect on $G_{Sref}$ relative to the direct effects of canopy architecture, and that height and sapwood-to-leaf area ratio explain most (75%) of the variation in $G_{Sref}$ among closed canopy ecosystems. To my knowledge, only one other attempt was made to derive a generic formulation for reference conductance, in which total canopy conductance at a reference state (i.e. $G_{Tref}$) was related to LAI (Granier et al., 2000). In that study, which considered a wide range of forested ecosystems ($n = 18$), $G_{Tref}$ increased linearly with LAI, saturating at about the midpoint of the LAI range. Here, the observed relationship between $G_{Sref}$ and LAI ($r^2=0.10$) is much weaker than the observed relationship of $G_{Sref}$ to $A_S/A_L/h$ ($r^2 = 0.84$) proposed here.

For this parsimonious formulation to have prognostic utility at coarse spatial scales, $A_S/A_L$ must be specified. At the ecosystem scale, this hydraulic characteristic is relatively simple to estimate when compared to the effort required to collect eddy covariance or sap flux data and the suite of meteorological measurements typically required to estimate $G_{Sref}$ at single stand. At the landscape scale, sapwood area may be estimated for monospecific stands with well-established allometric relationships with height or basal area measurements (Meinzer et al., 2005), both of which can be derived with reasonable accuracy from LiDAR measurements (Kellndorfer et al., 2004; Lefsky et al., 1999, 2002). However, I do not at this time know of a generic, prognostic model for $A_S/A_L$ that would facilitate the application of Equation 3.8 over coarse spatial scales (i.e. regional), though these results suggest climatic mediation of the relationship between $A_S/A_L$ and $h$ that could motivate future research. Finally, a strong relationship between $G_{Sref}$ and $h$ within temperate forests was elucidated and could be more immediately useful in coarse scale modeling efforts.
4

A characterization of the variability in ecosystem-scale model parameters using Bayesian inversion of eddy covariance data

4.1 Introduction

About half of the annual anthropogenic emissions of CO$_2$ are cycled through the terrestrial biosphere (Joos, 2003), and terrestrial ecosystems are thought to be a main driver of inter-annual variability in atmospheric CO$_2$ (Peylin et al., 2005; Rodenbeck et al., 2003). Hence, predicting current and future climate states requires a robust quantification of the net ecosystem exchange of CO$_2$ ($NEE$) and the closely coupled transpiration ($T$) flux between the atmosphere and the terrestrial biosphere. Because these fluxes have been directly observed over just a small fraction of the global land surface, models for the processes that drive these exchanges are needed for assessing carbon and water cycle dynamics over large areas. The models (i.e. DGVMs and stand-scale ecosystem models) are simplified representations of a large number of non-linear and complex processes, including photosynthetic carbon assimilation, transpiration, autotrophic and heterotrophic respiration, carbon allocation
and stand growth, and nutrient cycling (Cramer et al., 2001; Prihodko et al., 2008; Siqueira et al., 2006). These models have been shown to perform well under some scenarios (Friend et al., 2007), but much uncertainty remains, originating in part from parameter uncertainty (Jung et al., 2007; Tang and Zhuang, 2008). It is also worth noting that many model operate on a daily time step (Table 1.2), obscuring flux variation over diurnal time scales which is known to be significant (Siqueira et al., 2006).

A key challenge in ecosystem modeling at spatial scales well in excess of a stand level is the \textit{a priori} specification of model parameters. Typically, the parameters in large-scale terrestrial ecosystem models are assigned according to biome or plant functional type (PFT) based on measurements from a few species or biomes within that PFT (Sellars et al., 2006, Wang et al., 2007, Zaehle et al. 2005). However, key model parameters have been shown to vary widely within PFTs (for example, see Chapter 3 and Wullschleger 1993). Some parameters have even been shown to vary significantly among stands of the same species (Alsheimer et al., 1998; Delzon et al., 2004; Wilson et al., 2000), and even among individual trees of the same stand (Ryan et al., 2000; Schäfer et al., 2000). Significant inter-annual variation in key model parameters has also been documented (Chen et al., 2008). The static, PFT-based approach to model parameterization is further challenged by the fact that these parameters are often determined from measurements made at the leaf or tree scale (Wang et al., 2007), and simply incorporating these measurements into ecosystem scale models may ignore important emergent processes resulting from leaf-to-canopy up-scaling (Field et al., 1995; Leuning et al., 1995). Given the large uncertainty in the magnitude of biosphere-atmosphere CO$_2$ fluxes (Friend et al., 2007), improving parameterization schemes for coarse-scale ecosystem models remains a research priority.

Eddy covariance records of NEE and evapotranspiration (ET), which now exist
for over 400 ecosystems from a broad range of global PFTs (Baldocchi, 2008), form an unprecedented dataset for beginning to explore how to constrain parameters of DGVMs (Friend et al., 2007). Through standard optimization techniques such as cost function minimization or through Bayesian inference, these flux records may be inverted against models to determine the parameters of interest. Because the FLUXNET network of eddy covariance towers spans the full range of global biomes, the influence of climate, plant functional type, and structural properties of the ecosystem on the parameters may be readily assessed. Furthermore, the eddy covariance method permits the observation of ecosystem-scale processes directly, minimizing some of the challenges of up-scaling models from the leaf to canopy. Finally, flux records from many sites have been measured continuously for a number of years, permitting analysis of how model parameters may change across a number of growing seasons.

Eddy covariance records also have their limitations, including significant measurement error (Richardson et al., 2006), large and frequent gaps created by the rejection of data collected during stable atmospheric conditions, and an absence of data from highly heterogeneous site and sites with complex topography (Aubinet, 2008). Further, because eddy covariance measurements are made directly at the ecosystem scale, it can be difficult to link trends in the flux estimates to certain physical and biological process that vary over smaller spatial scales within the canopy or that require measured fluxes to be partitioned into primary components (like assimilation and respiration in the case of NEE). Nonetheless, the network of eddy covariance sites has produced one of the most diverse and useful datasets for identifying convergent trends in flux responses across ecosystems.

The objective of this study are to eddy covariance data explore the spatial and temporal variability NEE and ET model parameters with the goal of improving coarse-scale ecosystem models by constraining this variability with other observable
ecosystem attributes. To achieve this objective, data records from over 35 FLUXNET sites within the Ameriflux network in the eastern United States are inverted to determine model parameters for some of the simplest yet most widely used models for describing CO₂ and H₂O exchange between the biosphere and the atmosphere. These model parameters were then used to answer the following questions regarding parameter variability and model parameterization schemes:

1. **How much do these parameters vary across sites?**

2. **Does grouping the parameters according to PFT reduce uncertainty in model specification?**

3. **How much parameter variability can be attributed to inter-annual variation?**

4. **Do significant relationships exist between model parameters and climatic variables or vegetation indices that are readily available at coarse spatial scales? And if so, could these relationships be used to constrain these parameters?**

As I address the final question, I make an effort to explain observed cross-site correlation between model parameters and site indices using established physiological principles to facilitate the appropriate incorporation of these prognostic relationships in future modeling efforts.

Bayesian inference was used to determine parameters from the flux records using a Markov-Chain Monte Carlo Gibbs sampling technique commonly known as the Metropolis-Hastings (M-H) algorithm (Hastings, 1970; Metropolis et al., 1953). While a Bayesian framework may not necessarily outperform other optimization strategies in terms of finding the ‘correct’ parameter values (Trudinger et al., 2007), this approach does offer a number of other advantages over standard optimization techniques. First, the M-H approach produces full probability distribution functions
(PDFs) for the inverted model parameters, whereas deterministic approaches produce only point estimates. Analysis of the PDFs permits a characterization of the uncertainty of model parameters (Tang and Zhuang, 2008) and allows for a direct analysis of cross-correlation among the parameters themselves. Additionally, the Bayesian inference approach formally permits the incorporation of a priori information about the distribution of the model parameters through Bayes theorem, which can be expressed as:

$$P(\Theta|DS) \sim L(DS|\Theta)P(\Theta),$$

where $P(\Theta|DS)$ is the posterior probability distribution function for the set of parameters ($\Theta$) given the dataset ($DS$), $L(DS|\Theta)$ is the likelihood of $DS$ given $\Theta$, and $P(\Theta)$ is the prior probability distribution function for $\Theta$. Here, the prior distributions are non-informative uniform distributions with ranges that encompass values of the model parameters that have been previously reported in the literature. Future studies may be able to formally incorporate the form of the posterior distributions described here in efforts to reduce model uncertainty.

4.2 Model selection

Most DGVMs have a large number of parameters that are typically assigned on the basis of PFT (see Table 1.2). Recent work has shown that most of these models are over-parameterized (Prihodko et al., 2008; Tang and Zhuang, 2008), and significant correlations do exist between the parameters that lead to equifinality, or the tendency for achieving similar model accuracy with greatly different parameter pools. Furthermore, data assimilation studies that employ both standard and Bayesian optimization techniques suggest that only a limited number of parameters may be determined from eddy covariance data (Wang et al., 2001; Knorr and Kattge,
As all of these challenges are partially minimized by shrinking the manifold of parameter spaces, the focus here concerns low-dimensional models for $\text{NEE}$ and $T$ that have a small (i.e. $< 5$) number of parameters as a case study.

### 4.2.1 The non-rectangular hyperbola model for $\text{NEE}$

For $\text{NEE}$, I selected the widely used non-rectangular hyperbola model as proposed by Gilmanov et al. (2003):

$$\text{NEE} = \left[-\frac{1}{2\theta_f} \left( \alpha \text{PAR} + \beta - \sqrt{(\alpha \text{PAR} + \beta)^2 - 4\alpha\beta\theta_f \text{PAR}} \right) + \gamma \right]$$

where $\alpha$ is the mean apparent ecosystem quantum yield, $\beta$ is the light-saturated assimilation rate, $\theta_f$ describes the model curvature as the $\text{NEE-\text{PAR}}$ relationship approaches saturation, and the intercept parameter $\gamma$ - which is the flux rate in the absence of light – is an estimate of ecosystem respiration. In recent years, this model has been applied to eddy covariance data in a number of independent studies. However, those analyses were largely focused on estimating annual flux rates using site-specific calibrations (Gilmanov et al., 2010; Stoy et al., 2006b) or, in the case of Owen et al. (2007), focused on exploring relationships between the parameters themselves. Hence, the main novelty of this study is the exploration of inter-specific relationships between these parameters and information readily available at coarse spatial scales with the objective of improving coarse-scale model parameterization schemes.

### 4.2.2 A Jarvis-type model for canopy stomatal conductance

For transpiration, the common assumption of a negligible leaf-boundary layer resistance is invoked, in which case $T$ may be directly linked to canopy stomatal conductance per unit ground area ($G_T$) via:
\[ T = a_V \cdot G_T \cdot D \]  \hspace{1cm} (4.3)

I then selected a Jarvis-type model for stomatal conductance reformulated from Equation 2.2 to include a light reduction function as:

\[ G_T = G_{T_{\text{ref}}} \cdot f_1(D) \cdot f_2(PAR). \] \hspace{1cm} (4.4)

In this study, \( G_{T_{\text{ref}}} \) is defined for conditions of non-limiting light and soil moisture and a reference vapor pressure deficit of 1 kPa. The vapor pressure correction takes the form

\[ f_1(D) = \frac{b_d D^{-1/2} - 1}{b_d - 1}, \] \hspace{1cm} (4.5)

after Katul et al. (2009). An inverse logit function for the light response variable (i.e. \( f_2(PAR) = 1/(1 + e^{-c_1 PAR + c_2}) \), Avisser et al. 1985; Ward et al. 2008) was selected. A light response function is necessary to properly identify \( G_{T_{\text{ref}}} \) and \( b_d \); however, because \( c_1 \) and \( c_2 \) are empirical constants without strong physiological meaning, they will not be a primary focus of the analysis.

The Jarvis-type formulation is selected as opposed to other canopy conductance models for three reasons: a) in contrast with many leaf-level conductance models (Collatz et al., 1991; Leuning et al., 1995), this model is independent of measured or modeled assimilation fluxes. Therefore, a larger number of parameters may be determined from data inversion of the measured latent heat fluxes and carbon fluxes; b) the parameters of this model have previously been determined through Bayesian Gibbs sampling for at least one study (Ward et al., 2008), and c) theory describing expected patterns of canopy stomatal conductance across species have already been developed (see previous chapter for details).
The effect of soil moisture content on NEE and $G_T$ is not explicitly considered in this study due to the absence of soil moisture records for many FLUXNET sites which unfortunately limits the immediate application of results from this study to coarse scale modeling efforts. Precipitation data and long-term records of the Palmer Drought Severity Index (PDSI) obtained from NOAA are used to restrict the analysis to conditions when conductance is not likely to be water limited, as described in the methods section. Also, eddy covariance records measure $ET$, and not $T$. Therefore, $G_T$ data is further restricted to conditions when leaf area index (LAI) is greater than 2.5, and the time since a major rain event (arbitrarily defined as $> 5$ mm in a one hour period) is greater than 2 days. Under these conditions, I expect that soil and interception evaporation would be minimal and the $ET$ flux should be dominated by $T$.

4.2.3 Linking CO$_2$ and H$_2$O exchange

Transpiration and gross canopy assimilation ($A$) are closely-coupled fluxes and their relationship is often expressed in terms of water use efficiency, which may be defined as a flux-based water use efficiency ($\lambda_L = A/T$) or as a marginal water use efficiency ($\lambda = \partial A/\partial T$). Both $A$ and $T$ are commonly expressed with Fickian-diffusian type analogies as:

$$A = G_s/1.6(c_a - c_i)$$

and

$$T = G_s(e_i - e_a)$$

where $c_a$ and $c_i$ are the atmospheric and intercellular CO$_2$ concentrations, respectively, the constant 1.6 describes the relative diffusivity of water with respect to CO$_2$, and $e_i$ and $e_a$ are the intercellular and atmospheric vapor pressures, respectively. If it is assumed that $e_i$ is approximately equal to the saturation vapor pressure, then $D$
may be used a surrogate for $e_i - e_a$. With these formulation, the flux-based water use efficiency may be expressed as

$$\lambda_L = \frac{A}{T} = \frac{c_a}{1.6D} (1 - \frac{c_i}{c_a}),$$

which may be rearranged to give an expression for the ratio of intercellular to ambient CO$_2$ concentration:

$$\frac{c_i}{c_a} = 1 - \left( \frac{1.6\lambda_L}{c_a} \right) D.$$  

Another expression for $c_i/c_a$ dependent on the marginal water-use efficiency is presented in Katul et al. (2009):

$$\frac{c_i}{c_a} = 1 - \left( \frac{\lambda}{1.6/c_a} \right)^{1/2} D^{1/2}.$$  

Combining these two equation suggests that the flux-based water use efficiency varies on short (i.e. hourly) time scales as:

$$\lambda_L = \left( \frac{\lambda c_a}{1.6} \right)^{1/2} D^{-1/2}$$

As a final step in the analysis, short-term (i.e. hourly) variation in $\lambda_L$ with $D^{-1}$ was explored at each site. From the these relationships, $\lambda$ was estimated and compared across sites.

4.3 Methods

4.3.1 Eddy covariance and meteorological data – sources and quality control

The data used in this analysis are from a subset of FLUXNET sites in the Ameriflux regional network representing a range of biomes including 4 grassland ecosystems, 6 agricultural sites, 13 deciduous forests, 10 evergreen forests, and 1 wetland ecosystem (Table 4.1). Hourly-averaged eddy covariance fluxes and meteorological data such
as \( PAR, D, T_a, \bar{u}, u^*, R_n \) were obtained for these sites from the data archive at the Carbon Dioxide Information Analysis Center (CDIAC) maintained by Oak Ridge National Laboratory. Generally, fluxes and meteorological records were obtained for years 2001-2005, with some exceptions for sites that were not active during that period but were active in other years.

The data files downloaded from CDIAC contain quality control flags indicating flux measurements that the principal investigators from each site have screened as unreliable. All such flagged data were removed from the records. Then, because quality control methods may vary significantly from site to site (see Chapter 2), I applied an additional \( u^* \) filter similar to the method described by Reichstein and co-workers (2005). Finally, fluxes were rejected if the half-hourly flux magnitude was outside the logical expectation window of -50 to 100 \( \mu \text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1} \) for \( \text{NEE} \) (where a positive flux denotes vegetative carbon assimilation from the atmosphere) and -300 to 1000 W \( \text{m}^{-2} \) for \( \text{LE} \). Data meeting all of these quality control requirements are henceforth termed ”acceptable” flux measurements. I used flux estimates that were uncorrected for storage fluxes as the analysis is performed on fluxes collected during daylight hours when storage terms are negligible, and not all sites included profiling systems necessary for an accurate accounting of storage terms.

4.3.2 Ancillary data

\( \text{LAI} \) measurements from some sites were available from CDIAC at a high temporal resolution (i.e. monthly or weekly), and continuous \( \text{LAI} \) records for those sites were obtained by simply interpolating those measurements. In some cases, high resolution \( \text{LAI} \) data was not available from CDIAC but was available in the literature or from personal communication with principal investigators. For many forest sites, only minimum and maximum \( \text{LAI} \) values were available. In those cases, a continuous \( \text{LAI} \) record was obtained by scaling the smoothed, gap-filled MODIS 8-day \( \text{LAI} \)
<table>
<thead>
<tr>
<th>ID</th>
<th>Site</th>
<th>State</th>
<th>Lat</th>
<th>Lon</th>
<th>h</th>
<th>LAI</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wetlands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>001</td>
<td>Lost Creek</td>
<td>WI</td>
<td>48.08</td>
<td>-89.98</td>
<td>2</td>
<td>4.0</td>
<td>2002</td>
<td>Wang et al. (2006)</td>
</tr>
<tr>
<td>101</td>
<td>Duke Wetlands</td>
<td>NC</td>
<td>39.97</td>
<td>-79.09</td>
<td>0.7</td>
<td>1.4</td>
<td>2004</td>
<td>Novick et al. (2004)</td>
</tr>
<tr>
<td>102</td>
<td>Fermie Prairie</td>
<td>IL</td>
<td>41.84</td>
<td>-88.24</td>
<td>0.5</td>
<td>2.9</td>
<td>2006</td>
<td>Coulter et al. (2006)</td>
</tr>
<tr>
<td>103</td>
<td>Shindler</td>
<td>OK</td>
<td>36.93</td>
<td>-96.68</td>
<td>0.6</td>
<td>2.8</td>
<td>1999</td>
<td>Suyker et al. (2003)</td>
</tr>
<tr>
<td>104</td>
<td>Walnut River</td>
<td>KS</td>
<td>37.52</td>
<td>-96.86</td>
<td>0.5</td>
<td>2.8</td>
<td>2002</td>
<td>Coulter et al. (2006)</td>
</tr>
<tr>
<td>105</td>
<td>Shindler</td>
<td>OK</td>
<td>36.93</td>
<td>-96.68</td>
<td>0.6</td>
<td>2.8</td>
<td>1999</td>
<td>Suyker et al. (2003)</td>
</tr>
<tr>
<td>106</td>
<td>Walnut River</td>
<td>KS</td>
<td>37.52</td>
<td>-96.86</td>
<td>0.5</td>
<td>2.8</td>
<td>2002</td>
<td>Coulter et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Agricultural Sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>201</td>
<td>AR Mead</td>
<td>OK</td>
<td>36.61</td>
<td>-97.49</td>
<td>0.5</td>
<td>1.6</td>
<td>2004</td>
<td>Fischer et al. (2007)</td>
</tr>
<tr>
<td>202</td>
<td>Bondville Main</td>
<td>IL</td>
<td>40.01</td>
<td>-88.29</td>
<td>3/0.6</td>
<td>5.7 / 6.5 2003 (Maize) / 2004 (Soy)</td>
<td>Meyers and Hollinger (2004)</td>
<td></td>
</tr>
<tr>
<td>203</td>
<td>Fermie Agricultural</td>
<td>IL</td>
<td>40.01</td>
<td>-88.29</td>
<td>3/0.6</td>
<td>5.7 / 6.5 2003 (Maize) / 2004 (Soy)</td>
<td>Meyers and Hollinger (2004)</td>
<td></td>
</tr>
<tr>
<td>204</td>
<td>Fermie Agricultural</td>
<td>IL</td>
<td>40.01</td>
<td>-88.29</td>
<td>3/0.6</td>
<td>5.7 / 6.5 2003 (Maize) / 2004 (Soy)</td>
<td>Meyers and Hollinger (2004)</td>
<td></td>
</tr>
<tr>
<td>205</td>
<td>Fermie Agricultural</td>
<td>IL</td>
<td>40.01</td>
<td>-88.29</td>
<td>3/0.6</td>
<td>5.7 / 6.5 2003 (Maize) / 2004 (Soy)</td>
<td>Meyers and Hollinger (2004)</td>
<td></td>
</tr>
<tr>
<td>206</td>
<td>Fermie Agricultural</td>
<td>IL</td>
<td>40.01</td>
<td>-88.29</td>
<td>3/0.6</td>
<td>5.7 / 6.5 2003 (Maize) / 2004 (Soy)</td>
<td>Meyers and Hollinger (2004)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deciduous Forests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>301</td>
<td>Bartlett</td>
<td>NH</td>
<td>44.06</td>
<td>-71.29</td>
<td>19</td>
<td>6.0</td>
<td>2005</td>
<td>Jenkins et al. (2007)</td>
</tr>
<tr>
<td>302</td>
<td>Duke Hardwood</td>
<td>NC</td>
<td>35.14</td>
<td>-79.10</td>
<td>27</td>
<td>6.5</td>
<td>2004</td>
<td>Stoy et al. (2006b)</td>
</tr>
<tr>
<td>303</td>
<td>Harvard Forest</td>
<td>MA</td>
<td>42.54</td>
<td>-72.17</td>
<td>23</td>
<td>4.9</td>
<td>2004</td>
<td>Urbanski et al. (2007)</td>
</tr>
<tr>
<td>304</td>
<td>Little Prospect Hill</td>
<td>MA</td>
<td>42.54</td>
<td>-72.17</td>
<td>23</td>
<td>4.9</td>
<td>2004</td>
<td>Urbanski et al. (2007)</td>
</tr>
<tr>
<td>305</td>
<td>Morgan Monroe</td>
<td>IN</td>
<td>33.32</td>
<td>-86.41</td>
<td>27</td>
<td>4.9</td>
<td>2002</td>
<td>Schmid et al. (2000)</td>
</tr>
<tr>
<td>306</td>
<td>Missouri Ozark</td>
<td>MO</td>
<td>38.74</td>
<td>-92.20</td>
<td>17</td>
<td>3.9</td>
<td>2004</td>
<td>Gu et al. (2006)</td>
</tr>
<tr>
<td>307</td>
<td>Ohio Oak Openings</td>
<td>OH</td>
<td>41.33</td>
<td>-83.51</td>
<td>25</td>
<td>4.7</td>
<td>2004</td>
<td>Noormets et al. (2008b)</td>
</tr>
<tr>
<td>308</td>
<td>Sylvania</td>
<td>MI</td>
<td>46.24</td>
<td>-89.35</td>
<td>22</td>
<td>4</td>
<td>2002</td>
<td>Desai et al. (2005)</td>
</tr>
<tr>
<td>309</td>
<td>UMBS</td>
<td>MI</td>
<td>45.56</td>
<td>-84.71</td>
<td>21</td>
<td>6.7</td>
<td>2003</td>
<td>Gough et al. (2008)</td>
</tr>
<tr>
<td>310</td>
<td>Walker Branch</td>
<td>TN</td>
<td>36.96</td>
<td>-84.29</td>
<td>25</td>
<td>5.6</td>
<td>2004</td>
<td>Babcock and Wilson (2001)</td>
</tr>
<tr>
<td>311</td>
<td>Willow Creek</td>
<td>WI</td>
<td>45.81</td>
<td>-90.08</td>
<td>24</td>
<td>5.4</td>
<td>2002</td>
<td>Cook et al. (2004)</td>
</tr>
<tr>
<td>312</td>
<td>Wisconsin intermediate</td>
<td>WI</td>
<td>46.01</td>
<td>-91.23</td>
<td>6</td>
<td>2.9</td>
<td>2003</td>
<td>Ewers et al. (2008)</td>
</tr>
<tr>
<td>313</td>
<td>Wisconsin mature hardwood</td>
<td>WI</td>
<td>46.63</td>
<td>-91.10</td>
<td>21</td>
<td>3.8</td>
<td>2004</td>
<td>Noormets et al. (2008a)</td>
</tr>
<tr>
<td></td>
<td>Evergreen Forests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>401</td>
<td>Austin Cary</td>
<td>FL</td>
<td>29.74</td>
<td>-82.22</td>
<td>22</td>
<td>2.8</td>
<td>2004</td>
<td>Powell et al. (2008)</td>
</tr>
<tr>
<td>402</td>
<td>Donalson</td>
<td>FL</td>
<td>29.75</td>
<td>-82.16</td>
<td>13</td>
<td>3.3</td>
<td>2003</td>
<td>Gholz and Clark (2002)</td>
</tr>
<tr>
<td>403</td>
<td>Duke Pine</td>
<td>NC</td>
<td>35.98</td>
<td>-79.09</td>
<td>18</td>
<td>4.3</td>
<td>2004</td>
<td>Stoy et al. (2006b)</td>
</tr>
<tr>
<td>404</td>
<td>Howland</td>
<td>ME</td>
<td>45.20</td>
<td>-68.74</td>
<td>20</td>
<td>5.7</td>
<td>2004</td>
<td>Hollinger et al. (2004)</td>
</tr>
<tr>
<td>405</td>
<td>Kennedy Scrub Oak</td>
<td>FL</td>
<td>28.61</td>
<td>-80.67</td>
<td>2</td>
<td>1.7</td>
<td>2004</td>
<td>Bracho et al. (2008)</td>
</tr>
<tr>
<td>407</td>
<td>Mine</td>
<td>FL</td>
<td>29.77</td>
<td>-82.24</td>
<td>5</td>
<td>3.3</td>
<td>2004</td>
<td>Gholz and Clark (2002)</td>
</tr>
<tr>
<td>408</td>
<td>NC Lobolly Pine</td>
<td>NC</td>
<td>35.80</td>
<td>-76.67</td>
<td>14</td>
<td>3.9</td>
<td>2006</td>
<td>Noormets et al. (2010)</td>
</tr>
<tr>
<td>409</td>
<td>Wisconsin Mature Red Pine</td>
<td>WI</td>
<td>46.74</td>
<td>91.17</td>
<td>18</td>
<td>2.8</td>
<td>2004</td>
<td>Noormets et al. (2008b)</td>
</tr>
<tr>
<td>410</td>
<td>Wisconsin Young Pine</td>
<td>WI</td>
<td>46.62</td>
<td>-91.10</td>
<td>6</td>
<td>0.9</td>
<td>2004</td>
<td>Noormets et al. (2007)</td>
</tr>
</tbody>
</table>

Table 4.1: The Ameriflux study sites, grouped by plant functional type. The variable $h$ is canopy height in meters. 'Lat' and 'Lon' are latitude and longitude, respectively. The variable LAI refers to the average leaf area index during the study period.
records (Gao et al., 2008) by the minimum and maximum ground-based LAI measurements. MODIS LAI products have previously been shown to satisfactory reproduce temporal LAI patterns for a range of ecosystem types (Cohen et al., 2003; Fensholt et al., 2004; Wang et al., 2004). Typically, the MODIS product tends to overpredict LAI, though the scaling of the MODIS records by ground-based measurements should remove much of this bias in this study. Estimates of the Palmer Drought Severity Index (PDSI) were obtained from NOAA’s National Climatic Data Center (NCDC) for each site-year in the study domain. Estimates of long term mean annual temperature and precipitation, when they weren’t available from CDIAC, were also obtained from NOAA.

4.3.3 Additional data filtering and study period definition

In general, data inversions were performed using data from a single growing season, which was selected as the year 2004 growing season. Other years were chosen for some sites that lacked complete records of flux or meteorological variables, or which experienced a drought in 2004. In a few cases, other years were also selected to match the availability of a high-quality LAI record.

Additional filters beyond the quality control filters described in Section 4.3.1 were applied to the datasets to match climatic conditions to model assumptions, and to limit variability associated with changing LAI. Specifically, for the NEE inversions, data were limited to acceptable fluxes collected during daytime periods (i.e. PAR > 10 µmol m$^{-2}$ s$^{-1}$) when LAI was within 90% of the maximum value achieved during the study period and when the Palmer Drought Severity Index was greater than -1 (indicating near normal to wet conditions). These same filters were applied to the data used in the $G_T$ inversions, which were further limited to data collected more than two days after a major rain event when LAI > 2.5. There additional filters on the $G_T$ data were applied in an effort to ensure that measured latent heat fluxes were
not significantly ‘contaminated’ by interception or soil evaporation, which should be low in dry, dense canopies. Because the filters for the conductance inversions are more stringent, and because the Jarvis model requires an additional meteorological variable (i.e. $D$) that is not available for all sites, the parameters of the conductance model could not be determined for all sites in Table 4.1.

4.3.4 Error in the eddy covariance measurements

Random measurement error for $NEE$, $LE$, and $GT$ was determined using the paired-measurement technique recommended by Richardson et al. (2006). Specifically, the mean absolute error $\delta$ was determined by comparing acceptable flux measurements collected during the same time period on consecutive days under similar environmental conditions (i.e. wind speed within 1 m s$^{-1}$, $Ta$ within 3 °C, and $Rn$ within 75 W m$^{-2}$). The random measurement error was then calculated as the standard deviation of the absolute error ($\sigma(\delta)$) for consistency with the Richardson study.

4.3.5 Bayesian inversion

The $M$-$H$ algorithm within a Gibbs sampler was used to estimate the posterior probabilities of the inverted parameters (i.e. the $NEE$ parameters $\alpha$, $\beta$, $\gamma$, and $\theta_f$ and the conductance parameters $GT_{ref}$, $b_d$, $c_1$, and $c_2$) for the sites in Table 4.1. I began by assuming that the observations were independently and identically distributed and followed a normal distribution with standard deviations of $\varepsilon_{NEE}$ for $NEE$ and $\varepsilon_{GT}$ for $GT$. While previous work has suggested that random error in eddy covariance measurements is more closely approximated with a double-exponential distribution then a normal distribution (Richardson et al., 2006), other work suggests that error magnitude scales with flux magnitude, and that a normal distribution may be more appropriate when fluxes are limited to certain ranges in magnitude (Lasslop et al., 2008). The analysis is limited to daytime fluxes from the peak of the growing season.
when assimilation rates are expected to be quite high. Therefore, the distribution of errors in this analysis should be more Gaussian than the distribution of errors from data collected diurnally throughout the year. With these assumptions, the likelihood of Equation 4.1 becomes (Ricciuto et al., 2008):

\[
L(DS|\Theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\varepsilon^2}} \exp \left( -\frac{1}{2} \left[ \frac{f(\Theta) - DS_i}{\varepsilon} \right] \right)
\] (4.12)

where \( n \) is the length of the data record, \( f(\Theta) \) is the model (i.e. Eq. 4.2 for \( \text{NEE} \) and Eq. 4.5 for \( \text{GT} \)), and \( DS_i \) is the measured flux rate at time \( t_i \). The error term \( (\varepsilon) \) may be used to formally link the measured \( \text{NEE} \) and \( \text{GT} \) to the models via \( \text{NEE}_{\text{MEASURED}} = N(\text{NEE}_{\text{MODELED}}, \varepsilon_{\text{NEE}}) \), and \( \text{GT}_{\text{MEASURED}} = N(\text{GT}_{\text{MODELED}}, \varepsilon_{\text{GT}}) \), with

\[
\text{NEE}_{\text{MEASURED}} = N \left[ -\frac{1}{2\theta_f} \left( \alpha\text{PAR} + \beta - \sqrt{(\alpha\text{PAR} + \beta)^2 - 4\alpha\beta\theta_f\text{PAR}} \right) + \gamma, \varepsilon_{\text{NEE}} \right]
\] (4.13)

and

\[
\text{GT}_{\text{MEASURED}} = N \left[ \text{GT}_{\text{REF}} \left( \frac{b_dD^{-1/2} - 1}{d^b - 1} \right) \left( \frac{1}{1 + e^{-\alpha_1\text{PAR} + \alpha_2}} \right), \varepsilon_{\text{GT}} \right].
\] (4.14)

The prior distributions for the Bayesian inversions were assumed to be non-informative uniform distributions with the very wide ranges given in Table 4.2, permitting the posterior marginal distributions to be primarily determined by the data.

The \( \text{NEE} \) and \( \text{GT} \) model parameters are independent and the inversions were performed with two independent Gibbs samplers using methods outlined in Clark (2007). In both cases, the parameters were sampled using a multivariate normal jump distribution that assumed no correlation among the parameters. The covariance matrix of the jump distribution was varied from site to site to approach a target
Table 4.2: The range of the uniform prior distributions for the model parameters. Units are mg CO$_2$ µmol$^{-1}$ for the apparent ecosystems quantum yield ($\alpha$), mg CO$_2$ m$^{-2}$ s$^{-1}$ for the maximum ecosystem assimilation ($\alpha$) and the ecosystem respiration parameter ($\gamma$), mmol m$^{-2}$ s$^{-1}$ for the total reference canopy conductance ($G_{Tref}$), and kPa$^{-1/2}$ for the stomatal sensitivity parameter ($b_d$). $c_1$ (units of m$^2$ s µmol$^{-2}$) and $c_2$ (dimensionless) are the two light response parameters. The curvature parameter $\theta_f$ is dimensionless.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NEE model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.1</td>
<td>1.5</td>
</tr>
<tr>
<td>$\theta_f$</td>
<td>0.1</td>
<td>3</td>
</tr>
<tr>
<td><strong>$G_T$ model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_{Tref}$</td>
<td>0</td>
<td>2000</td>
</tr>
<tr>
<td>$b_d$</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>$c_1$</td>
<td>0</td>
<td>$5 \times 10^{-3}$</td>
</tr>
<tr>
<td>$c_2$</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

acceptance rate of 0.5 to 0.7 for the proposed parameter values. The posterior distributions for $\varepsilon_{NEE}$ and $\varepsilon_{GT}$ were sampled directly in a separate step of the $M-H$ algorithm as described in Clark (2007). The Gibbs samplers were run for 120,000 iterations, and the parameter chains were ‘thinned’ by extracting every 200th sample. The first half of the thinned chains were then discarded to remove burn-in periods, leaving 300 parameter estimates to build the posterior PDFs.

The $M-H$ algorithm was used to generate one set of parameter values for each site in Table 4.1, with two exceptions. First, for managed agricultural sites with corn-soy rotations (i.e. Sites 202, 204 and 205), the algorithm was run twice: once using data from a year in which corn was planted, and once using data from a year in which soy was planted. Second, for a subset of sites (Sites 101, 202, 305, 308, 403, and 407), the algorithm was run three times using three individual years of data to
permit an analysis of inter-annual variability in model parameters.

4.3.6 Determining water use efficiency

As a final step in the analysis, hourly $\lambda_L$ was estimated for each site. To determine $A$ (necessary for calculating $\lambda_L$), I subtracted the average value of the respiration parameter $\gamma$ derived from the posterior PDFs from the measured carbon fluxes. I note that Eq. 4.2 is frequently used to partition eddy covariance carbon fluxes into ecosystem $A$ and respiration (Reichstein et al., 2005). The the respiration parameter may be treated as a constant for time periods on the order of weeks (Stoy et al., 2006b), which is close to the average length of the study periods in this analysis.

Next, the marginal water use efficiency $\lambda$, which can be assumed to be constant at least for short time periods (Katul et al., 2009), was estimated for each site using non-linear regression techniques from the relationship between $\lambda_L$ and $D^{-1/2}$.

4.4 Results

4.4.1 Data availability

The application of the eddy covariance data quality control procedures described in Section 4.3.1 removed a variable proportion of the available data across the sites in Table 4.1, with as much as 70% of the data record remaining in some sites and as little as 25% remaining in others. The LAI and PDSI filters described in Section 4.3.3 limited analysis to non-drought periods during the peak of the growing season and significantly shortened the data records; nonetheless, at most sites, 100-400 daytime hourly flux estimates remained for use in the NEE inversions. As described in the methods section, additional filters were applied to the data for use in the $G_T$ inversions, including the removal of data collected within 2 days of a rain event or when LAI < 2.5, and the removal for data for which an estimate of $D$ was unavailable. These additional filters further reduced the number of hourly flux estimates available
for the inversions and made these simulations impossible for some sites (e.g. Sites 201, 202, 206, 304, 310, 405, 406, and 410). The relationship between $G_T$ and $c_t/c_a$ was assessed for an even smaller number of sites due to the requirement of a continuous record of $c_a$, which was not universally available.

### 4.4.2 Random measurement error

The average random measurement error for $NEE$ measurements($\sigma(\delta)_{NEE}$), conductance measurements ($\sigma(\delta)_{G_s}$), and latent heat flux measurements ($\sigma(\delta)_{LE}$) was determined using the paired-point technique described in Section 4.3.4. $PFT$-averaged values of these error terms are given in Table 4.3. Across sites, $\sigma(\delta)_{NEE}$ averaged 0.15 mg CO$_2$ m$^{-2}$ s$^{-1}$ (or 3.4 µmol CO$_2$ m$^{-2}$ s$^{-1}$), which is comparable to, though slightly higher than, the mean measurement error of $\langle\sigma(\delta)_{NEE}\rangle = 2.6$ µmol m$^{-2}$ s$^{-1}$ reported in a previous study (Richardson et al., 2006) for growing season $NEE$. However, the data used in this analysis are limited to daytime conditions, which are associated with higher errors than the diurnal measurements analyzed in the Richardson study. Error in $NEE$ did not vary significantly across $PFT$s ($\langle(\sigma(\delta)_{NEE}\rangle = 0.15, 0.16, 0.16,$ and 0.15 mg CO$_2$ m$^{-2}$ s$^{-1}$ for grasslands, agricultural sites, deciduous forests, and evergreen forests, respectively).

The average random measurement error for latent heat flux was 41 W m$^{-2}$ across the entire dataset. This is well within the range of error in daytime $LE$ fluxes (21 to 170 W m$^{-2}$) reported in the Richardson study. Average $\langle\sigma(\delta)_{LE}\rangle$ is highest in deciduous forest ecosystems ($\langle\sigma(\delta)_{LE}\rangle = 43.6$ W m$^{-2}$), and in the grassland ecosystems ($\langle\sigma(\delta)_{LE}\rangle = 32.9$ W m$^{-2}$). Random measurement error for $G_T$ measurements averaged 127 mmol m$^{-2}$ s$^{-1}$ across the entire dataset. $\langle\sigma(\delta)_{G_T}\rangle$ was also high for deciduous forest ($\langle\sigma(\delta)_{G_T}\rangle = 210$ mmol m$^{-2}$ s$^{-1}$) and was lowest for agricultural sites ($\langle\sigma(\delta)_{G_T}\rangle =105$ mmol m$^{-2}$ s$^{-1}$). Across all sites, the error in $LE$ and the error in $G_T$ were not significantly correlated.
Table 4.3: The PFT-averaged model-data errors for the NEE and GT models (i.e. \( \varepsilon_{\text{NEE}}, \varepsilon_{\text{GT}} \)) estimated from the Bayesian inversions and the random measurement errors in NEE, LE, and GT (i.e. \( \sigma(\delta)_{\text{NEE}}, \sigma(\delta)_{\text{LE}}, \sigma(\delta)_{\text{GT}} \)) as estimated using the paired-measurement technique described in Section 4.3.4. The values in parentheses show the full range in error estimates within each PFT. Units are mg CO₂ m⁻² s⁻¹ for NEE, mmol m⁻² s⁻¹ for GT, and W m⁻² for LE.

<table>
<thead>
<tr>
<th>PFT</th>
<th>( \varepsilon_{\text{NEE}} )</th>
<th>( \sigma(\delta)_{\text{NEE}} )</th>
<th>( \varepsilon_{\text{GT}} )</th>
<th>( \sigma(\delta)_{\text{LE}} )</th>
<th>( \sigma(\delta)_{\text{GT}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wetland</td>
<td>0.12</td>
<td>0.08</td>
<td>53</td>
<td>44</td>
<td>32</td>
</tr>
<tr>
<td>Grasslands</td>
<td>.21 (.11 - .36)</td>
<td>.15 (.08 - .26)</td>
<td>98 (41 - 140)</td>
<td>33 (25 - 44)</td>
<td>129 (85 - 172)</td>
</tr>
<tr>
<td>Agricultural Sites</td>
<td>.21 (.11 - .26)</td>
<td>.16 (.08 - .21)</td>
<td>88 (55 - 162)</td>
<td>40 (28 - 61)</td>
<td>105 (24 - 309)</td>
</tr>
<tr>
<td>Deciduous Forests</td>
<td>.23 (.11 - .32)</td>
<td>.16 (.08 - .25)</td>
<td>217 (74 - 382)</td>
<td>44 (23 - 86)</td>
<td>211 (86 - 376)</td>
</tr>
<tr>
<td>Evergreen Forests</td>
<td>.20 (.10 - .32)</td>
<td>.15 (.07 - .21)</td>
<td>152 (82 - 247)</td>
<td>35 (20 - 48)</td>
<td>161 (99 - 224)</td>
</tr>
</tbody>
</table>

4.4.3 Bayesian inversion - posterior PDFs

The central moments of the posterior PDFs for the primary model parameters are presented in Figure 4.1, and means and 95% confidence intervals given in Table 4.4. In general, the posterior distributions were well constrained, with standard deviations approximately an order of magnitude less than the mean value of the distributions (Figure 4.1). The curvature parameter \( \theta_f \) was the most poorly constrained parameter, with average coefficients of variation (C.V.) for the posterior distributions of 46%. This finding is not unexpected as this parameter reflects the up-scaled effects of leaf-level processes on photosynthesis and hence is likely to be sensitive to a number of parameters including vertical variations in leaf area density distribution, sun angle, and air temperature. The posterior distributions of the conductance parameters were better constrained than NEE model posteriors, with average C.V.’s of \(<10\%\) for conductance model parameters, compared to C.V.’s of 0.17, 0.11, and 0.14 for \( \alpha \), \( \beta \), and \( \gamma \), respectively. The posterior PDF means for these parameters were within the range of parameter values previously reported in the literature (Gilmanov et al., 2003, 2010; Katul et al., 2009; Granier et al., 2000).

The posterior PDFs were close to, but not necessarily, Gaussian. The aver-
age skewness ($\gamma_1$) of the posterior distributions was near zero (Figure 4.1), though posterior distributions for many of the parameters were slightly skewed. In most cases, distributions were positively skewed though some posterior distributions for the parameter $\gamma$ were negatively skewed. The average kurtosis ($\gamma_2$) was near 3 for all parameters ($\langle \gamma_2 \rangle = 2.8, 3.1, 3.0, 2.9, 2.9, \text{ and } 2.9$ for $\alpha, \beta, \theta_f, \gamma, G_{Tref}, \text{ and } b_d$, respectively). However, some distributions are associated with significantly higher kurtosis values (Figure 4.1).

An example of the correlation among the parameter posteriors is shown in Figure 4.2 for the Duke Pine site (Site 403). Across sites, the quantum yield parameter $\alpha$ was generally correlated with $\beta$, $\gamma$, and $\theta_f$ ($\langle r^2 \rangle = 0.29, 0.58, \text{ and } 0.43$, respectively). The maximum assimilation rate $\beta$ was also generally correlated with $\theta_f$ ($\langle r^2 \rangle = 0.63$). Among the parameters of the conductance model, $G_{Tref}$ was generally correlated with the light response parameter $c_1$ ($\langle r^2 \rangle = 0.44$). Correlations among the other parameters was very low (i.e. $r^2 < 0.2$).

The posterior distributions of the model-data error terms $\varepsilon_{NEE}$ and $\varepsilon_{GT}$ were fairly well constrained, with average C.V.’s of .03 and .05, respectively. Average values within each PFT are given in Table 4.3 along with the random measurement error. Across sites, $\varepsilon_{NEE}$ was about 40% higher than the random measurement error for NEE and was not significantly higher within any PFT. Across sites, the average value of $\varepsilon_{GT}$ (122 mmol m$^{-2}$ s$^{-1}$) was comparable to the average value of $\sigma(\delta)_{GT}$ (127 mmol m$^{-2}$ s$^{-1}$).

The mean value of the posterior distributions was taken as the best estimate of the parameter value. Throughout the rest of this section and the discussion section, any references made to the parameters refer to the mean values of the posterior PDFs, unless otherwise noted.
<table>
<thead>
<tr>
<th>ID</th>
<th>$\alpha$ (0.5 - 0.7)</th>
<th>$\beta$ (0.7 - 1.0)</th>
<th>$\gamma$ (0.12 - 0.18)</th>
<th>$\theta_f$ (0.77 - 0.99)</th>
<th>$G_{Trsf}$</th>
<th>$b$</th>
<th>$c_1 \times 10^{-3}$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>0.6</td>
<td>0.8</td>
<td>0.15</td>
<td>0.93</td>
<td>0.518</td>
<td>26</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>002</td>
<td>2</td>
<td>2.2</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>003</td>
<td>1.6</td>
<td>1.6</td>
<td>0.18</td>
<td>0.38</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>004</td>
<td>1.4</td>
<td>1.2</td>
<td>0.22</td>
<td>0.27</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>005</td>
<td>1.7</td>
<td>1.2</td>
<td>0.28</td>
<td>0.72</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>006</td>
<td>1.5</td>
<td>1.2</td>
<td>0.11</td>
<td>0.24</td>
<td>0.72</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>007</td>
<td>1.8</td>
<td>1.2</td>
<td>0.38</td>
<td>0.72</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>008</td>
<td>2.0</td>
<td>2.2</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>009</td>
<td>1.9</td>
<td>1.5</td>
<td>0.18</td>
<td>0.72</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>010</td>
<td>1.4</td>
<td>1.4</td>
<td>0.22</td>
<td>0.27</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Wetlands

<table>
<thead>
<tr>
<th>ID</th>
<th>$\beta$ (0.7 - 1.0)</th>
<th>$\gamma$ (0.12 - 0.18)</th>
<th>$\theta_f$ (0.77 - 0.99)</th>
<th>$G_{Trsf}$</th>
<th>$b$</th>
<th>$c_1 \times 10^{-3}$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>010</td>
<td>1.0</td>
<td>0.9</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>011</td>
<td>1.8</td>
<td>1.6</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>012</td>
<td>1.6</td>
<td>1.4</td>
<td>0.18</td>
<td>0.38</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>013</td>
<td>1.4</td>
<td>1.2</td>
<td>0.22</td>
<td>0.27</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>014</td>
<td>1.7</td>
<td>1.4</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Grasslands

<table>
<thead>
<tr>
<th>ID</th>
<th>$\beta$ (0.7 - 1.0)</th>
<th>$\gamma$ (0.12 - 0.18)</th>
<th>$\theta_f$ (0.77 - 0.99)</th>
<th>$G_{Trsf}$</th>
<th>$b$</th>
<th>$c_1 \times 10^{-3}$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>015</td>
<td>1.0</td>
<td>0.9</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>016</td>
<td>1.8</td>
<td>1.6</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>017</td>
<td>1.6</td>
<td>1.4</td>
<td>0.18</td>
<td>0.38</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>018</td>
<td>1.4</td>
<td>1.2</td>
<td>0.22</td>
<td>0.27</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>019</td>
<td>1.7</td>
<td>1.4</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Agricultural Sites

<table>
<thead>
<tr>
<th>ID</th>
<th>$\beta$ (0.7 - 1.0)</th>
<th>$\gamma$ (0.12 - 0.18)</th>
<th>$\theta_f$ (0.77 - 0.99)</th>
<th>$G_{Trsf}$</th>
<th>$b$</th>
<th>$c_1 \times 10^{-3}$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>020</td>
<td>1.0</td>
<td>0.9</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>021</td>
<td>1.8</td>
<td>1.6</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>022</td>
<td>1.6</td>
<td>1.4</td>
<td>0.18</td>
<td>0.38</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>023</td>
<td>1.4</td>
<td>1.2</td>
<td>0.22</td>
<td>0.27</td>
<td>0.43</td>
<td>21</td>
<td>0.9</td>
</tr>
<tr>
<td>024</td>
<td>1.7</td>
<td>1.4</td>
<td>0.08</td>
<td>0.27</td>
<td>0.91</td>
<td>21</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Deciduous Forests

<table>
<thead>
<tr>
<th>ID</th>
<th>$\beta$ (0.7 - 1.0)</th>
<th>$\gamma$ (0.12 - 0.18)</th>
<th>$\theta_f$ (0.77 - 0.99)</th>
<th>$G_{Trsf}$</th>
<th>$b$</th>
<th>$c_1 \times 10^{-3}$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>030</td>
<td>1.8</td>
<td>1.2</td>
<td>0.10</td>
<td>0.08</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>031</td>
<td>1.5</td>
<td>1.0</td>
<td>0.16</td>
<td>0.11</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>032</td>
<td>1.8</td>
<td>1.4</td>
<td>0.18</td>
<td>0.14</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>033</td>
<td>2.6</td>
<td>2.1</td>
<td>0.15</td>
<td>0.15</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>034</td>
<td>2.5</td>
<td>2.0</td>
<td>0.17</td>
<td>0.11</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
</tbody>
</table>

### Evergreen Forests

<table>
<thead>
<tr>
<th>ID</th>
<th>$\beta$ (0.7 - 1.0)</th>
<th>$\gamma$ (0.12 - 0.18)</th>
<th>$\theta_f$ (0.77 - 0.99)</th>
<th>$G_{Trsf}$</th>
<th>$b$</th>
<th>$c_1 \times 10^{-3}$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>040</td>
<td>0.5</td>
<td>0.4</td>
<td>0.14</td>
<td>0.11</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>041</td>
<td>0.9</td>
<td>0.8</td>
<td>0.22</td>
<td>0.17</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>042</td>
<td>1.5</td>
<td>0.8</td>
<td>0.34</td>
<td>0.20</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>043</td>
<td>4.3</td>
<td>3.9</td>
<td>0.50</td>
<td>0.92</td>
<td>7.5</td>
<td>1.9</td>
<td>2.2</td>
</tr>
<tr>
<td>044</td>
<td>1.7</td>
<td>1.4</td>
<td>0.33</td>
<td>0.27</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>045</td>
<td>1.4</td>
<td>1.1</td>
<td>0.12</td>
<td>0.10</td>
<td>0.5</td>
<td>6.3</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Table 4.4: The means and the 95% confidence intervals on the posterior distributions. Units and symbols are as defined in Table 4.2.
4.4.4 Parameter variation across PFTs, within PFTs, and with time

Across all sites, parameter variation is large, with coefficients of variation ranging from 0.2 - 0.3 for $G_{Tref}$ and the light response parameters of the conductance model (i.e. $c_1$ and $c_2$), and up to 0.69 for $\beta$ (Table 4.5). On average, the cross-site variation in the $NEE$ model parameters ($\langle C.V. \rangle = 0.44$) was higher than the cross-site variation in $G_T$ model parameters ($\langle C.V. \rangle = 0.30$).

The $PFT$-averaged mean value for select model parameters, along with parameter ranges and 95% confidence intervals for the mean, are presented in Figure 4.3, where the confidence intervals were determined using a non-parametric bootstrap. For most parameters, the ranges overlap considerably across PFTs, and in many cases even the confidence interval for the mean overlap across PFTs. Notable exceptions are particularly high values of $\beta$ and $G_{Tref}$ in agricultural systems and low values of $\theta_f$ in grassland ecosystems.

In fact, variation within PFTs is only marginally lower than variation across PFTs. As the within-PFT coefficients of variation given in Table 4.5 illustrate, the average coefficient of variation is nearly as high within plant functional types ($\langle C.V. \rangle = .30$) as across plant functional types ($\langle C.V. \rangle = 0.37$). Or, in other words, grouping variables by PFT appears to do little to explain the variation in the model parameters.

To test how parameter distributions vary from year to year within a site, the inversions were performed with three years of data from a subset of the sites in Table 4.2 (Sites 101, 202, 305, 308, 403, and 407). Examples of these distributions from multiple years are shown in Figure 4.4 for the Duke Old Field (site 101, Figure 4.4a-f), the Bondville site (site 202, Figure 4.4g-l), and the Mize site (site 407, Figure 4.4m-r), and inter-annual $C.V.s$ averaged across all six sites are given in Table 4.5.
Figure 4.1: Histograms of the central moments of the posterior parameter distributions for the sites of Table 4.1. The distribution means are shown in the top row. Standard deviations ($\sigma$), skewness values ($\gamma_1$), and kurtosis values ($\gamma_2$) are shown in the second, third, and fourth rows, respectively.

For some parameters, the distributions were fairly similar from one year to the next (i.e. $\alpha$ for site 101 and $\beta$ for site 407, Figure 4.4a and 4.4n, respectively). However, in many cases the distributions varied significantly across years, with average inter-annual coefficients of variation ranging from $\sim 0.1$ for $G_{Tref}$ to $\sim 0.50$ for $\theta_f$, though inter-annual variation was generally lower than inter-PFT and intra-PFT variation, particularly for $\beta$ and $G_{Tref}$. 
4.4.5 Parameter variation across PFTs and with canopy architecture

The correlations between the model parameters and five structural and climatic indices - $h$, mean study period $LAI$, mean annual temperature ($T_M$), mean daytime growing season ($D_{AVG}^{1/2}$), and mean annual precipitation ($P_M$) were assessed. Relationships between $\beta/LAI$, $G_{Tref}/LAI$, and $\gamma/LAI$ and these indices were also investigated. Significant inter-specific relationships existed between most of the model parameters and at least one of the indices; however, in general the correlations were quite weak (i.e. $r^2 < 0.25$). Stronger correlations emerged between $\beta/LAI$, $\gamma/LAI$, and $G_{Tref}/LAI$ and $h$ ($r^2 = 0.39, 0.34$, and $0.46$, respectively) and these relation-
Figure 4.3: The mean parameter value within each PFT (thick black line) bounded by the 95% confidence intervals for the mean as determined from a non-parametric bootstrap (thin lines) and the full range of parameters within each PFT (dotted line). Ranges and 95% confidence intervals could not be determined for wetlands as only one wetland site was included in this study. Units for the parameters are as defined in Table 4.2.

Figure 4.5: The relationship between $\beta/LAI$ and $h$ improves when the wetland site is removed ($r^2 = 0.45$). The relationships between $\beta/LAI$ and $\gamma/LAI$ are also stronger among the non-managed sites, or forests ($r^2 = 0.46$ and $r^2 = 0.73$, respectively). Surprisingly, only weak relationships were observed between the un-normalized parameters $\beta$, $\gamma$, and $G_{Tref}$ and LAI directly (Figure 4.6).
Table 4.5: The coefficients of variation (C.V.s) of the model parameters. Inter-annual C.V.s were determined by averaging the C.V.s of posterior distributions from a subset of the sites from Table 4.1 for which the Bayesian inversions were run for multiple years of data. The inter-PFT C.V.s represent the variation in parameter values across all sites, and the remaining columns show C.V.s for all sites within each respective PFT. Symbols and units are defined in Table 4.2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Inter-annual</th>
<th>Inter-PFT</th>
<th>Grasslands</th>
<th>Agricultural sites</th>
<th>Deciduous forests</th>
<th>Evergreen forests</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.26</td>
<td>0.28</td>
<td>0.32</td>
<td>0.18</td>
<td>0.26</td>
<td>0.28</td>
<td>0.26</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.19</td>
<td>0.69</td>
<td>0.28</td>
<td>0.46</td>
<td>0.3</td>
<td>0.24</td>
<td>0.36</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.27</td>
<td>0.38</td>
<td>0.41</td>
<td>0.36</td>
<td>0.41</td>
<td>0.38</td>
<td>0.35</td>
</tr>
<tr>
<td>$\theta_f$</td>
<td>0.48</td>
<td>0.42</td>
<td>0.25</td>
<td>0.4</td>
<td>0.42</td>
<td>0.32</td>
<td>0.38</td>
</tr>
<tr>
<td>$G_{Tref}$</td>
<td>0.10</td>
<td>0.27</td>
<td>0.32</td>
<td>0.14</td>
<td>0.19</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>$b$</td>
<td>0.32</td>
<td>0.41</td>
<td>0.33</td>
<td>0.3</td>
<td>0.45</td>
<td>0.56</td>
<td>0.40</td>
</tr>
<tr>
<td>$c_1$</td>
<td>0.35</td>
<td>0.28</td>
<td>0.17</td>
<td>0.3</td>
<td>0.32</td>
<td>0.15</td>
<td>0.26</td>
</tr>
<tr>
<td>$c_2$</td>
<td>0.17</td>
<td>0.23</td>
<td>0.62</td>
<td>0.09</td>
<td>0.21</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.27</strong></td>
<td><strong>0.37</strong></td>
<td><strong>0.34</strong></td>
<td><strong>0.27</strong></td>
<td><strong>0.32</strong></td>
<td><strong>0.29</strong></td>
<td><strong>0.31</strong></td>
</tr>
</tbody>
</table>

4.4.6 The relationship between $\lambda_L$ and $D$

On average, $D^{-1/2}$ explained 35% of the variation in hourly $\lambda_L$ within each site, with much higher correlation ($r^2 > 0.60$) observed in some sites (for example, see Figure 4.7a). The marginal water use efficiency, assumed here to be constant for the duration of the study period, averaged $958 \pm 406 \mu$mol mol$^{-1}$ for C3 ecosystems, and averaged $2503 \pm 190 \mu$mol mol$^{-1}$ for C4 ecosystems (i.e. the maize-based agricultural sites). The coefficient of variation in $\lambda$ for all C3 sites was 0.41, with C.V.’s of 0.57, 0.37, 0.49, and 0.29 within grasslands, C3 agricultural sites, deciduous forests, and evergreen forests, respectively.

Vapor pressured deficit also explains a significant amount of variation in $\lambda$ across sites over longer time scales. Specifically, $D_{AVG}$ explains roughly 27% of the cross site variation in $\lambda$ when C4 sites are neglected from the regression (Figure 4.7b). The $\lambda$ for the Duke Old Field site (site 101) is associated with the greatest departure from the cross-site regression model. LAI in this ecosystem was estimated to average $2.7 \text{ m}^2 \text{ m}^{-2}$ during the study period based on above and below-canopy PAR measurements.
Figure 4.4: Inter-annual variability in the posterior distributions for select model parameters shown as relative frequency plots of the posterior PDFs generated using three separate years of data for Site 101 (a-f), Site 202 (g-l), and Site 407 (m-r). Parameters and units are defined in Table 4.2.

Reflecting light attenuation not only from green biomass, but also from dead biomass which comprises a significant portion of the total biomass from this site. Thus, while data were filtered to reject those collected when $LAI < 2.5 \text{ m}^2 \text{ m}^{-2}$ to ensure that the majority of measured $ET$ was $T$, in the Duke Old Field it is likely that this filter alone can not sufficiently filter out periods when $E$ contributed largely to total $ET$ and reduced the apparent $\lambda$. When this site is neglected from the regression, $D$ explains 38% of the cross-site variation in $\lambda$. 

111
Figure 4.5: The variation in (a) $G_{Tref}$, (b) $\beta$, and (c) $\gamma$ with canopy height ($h$). Figures (d) - (f) show the same relationships when the parameters are normalized by $LAI$. The regression line shown in (d) was determined using all the available data. In (e), the regression line was determined using parameters from all sites except the wetland, and in (f) the line was determined using forest ecosystems alone. Error bars show the 95% confidence intervals on the posterior distributions. Symbols are as follows: squares represent grassland ecosystems, circles represent agricultural sites, upward pointing triangles represent deciduous forests, right pointing triangles represent evergreen forests, and the diamond represents the wetland ecosystem. Units are as defined in Table 4.2

4.5 Discussion

Terrestrial ecosystem models are critical for quantifying and predicting biological components of the global carbon and water cycles. These models are complex, highly-nonlinear, and dependent on a large number of parameters which have been shown to inject a large amount of uncertainty in model outputs (Klemmedtsson et al., 2008; Tang and Zhuang, 2008). Typically, these parameters are assumed to be static quantities specified on the basis of $PFT$, despite well-documented variation of these parameters
within PFTs and with time. Furthermore, these parameters are often determined from measurements at scales which do not match the scales at which the models are applied (i.e. leaf vs. ecosystem scales).

In this study, I applied Bayesian inversion techniques to data from over 35 eddy covariance flux measurements sites in the eastern United States to estimate the parameters of low-dimensional models for vegetative CO$_2$ and H$_2$O exchange, and then used these parameter estimates to answer a series of questions as outlined in the introduction. First, I asked "How much do these parameters vary across PFTs?" As Tables 4.4 and 4.5 illustrate, the parameters vary significantly across sites. The maximum assimilation parameter $\beta$, for example, varies by nearly an order of magnitude from $\beta = 0.6$ mg CO$_2$ m$^{-2}$ s$^{-2}$ in the Wisconsin young pine site (site 410) to $\beta = 6.4$ mg CO$_2$ m$^{-2}$ s$^{-2}$ in the Mead irrigated Maize/Soy cropland (Site 204). Noticeably large cross-site variation was also observed for the curvature parameter $\theta_f$ (which ranged from 0.15 -0.95) and the stomatal sensitivity parameter $b_d$ (which ranged from 1.5 to 6.6 kPa$^{1/2}$). Overall, the average cross-site coefficient of variation for the model parameters was 0.37 (Table 4.5). Slightly higher variation was observed, on average, for the NEE model parameters ($\langle C.V. \rangle = 0.44$) as compared to the conductance model parameters ($\langle C.V. \rangle = 0.30$), which is not surprising given that net carbon exchange is a complex process that incorporates stomatal conductance via its link to assimilation (i.e. Eq. 4.6), but which also incorporates photosynthetic and respiration processes.

After determining the extent of the variation across all sites, I then asked "Does grouping the parameters according to PFT reduce uncertainty in parameter specification?" By examining the intra-PFT C.V.s (i.e. Table 4.5), as well as the mean parameter values and ranges shown in Figure 4.3, it becomes clear that grouping the parameters by plant functional type does not appear to be a good prognostic tool for specifying parameter values. The parameter C.V.s within PFTs are just slightly
Figure 4.6: The variation in (a) $G_{\text{Trf}}$, (b) $\beta$, and (c) $\gamma$ with leaf area index ($LAI$). Error bars show the 95% confidence intervals on the posterior distributions. Symbols are as in Figure 4.5, and units are as defined in Table 4.2.

lower than the cross-$PFT$ C.V.'s. The only $PFT$ grouping associated with noticeable reduction in parameter variation was agricultural sites, where the within-$PFT$ C.V. was less than the cross-$PFT$ C.V. for the majority of the parameters; however, even within these agricultural sites, variation remains high ($\langle C.V. \rangle = 0.9$ to 0.46).

Not only does a $PFT$ classification scheme fail to significantly reduce the degree of parameter variation, it also does little to discriminate average parameter values. The ranges of parameter values associated with each $PFT$ overlap considerably (Figure 4.3), as do, in many cases, the 95% confidence intervals for the $PFT$-averaged means. There are some exceptions; for example the noticeably low values of the curvature
Figure 4.7: The relationship between flux-based water use efficiency $\lambda_L$ and $D^{-1/2}$ is shown in (a) for site 204. The cross-site relationship between marginal water use efficiency ($\lambda$) long-term averaged $D$ is shown in (b) for all sites for which $\lambda$ could be determined. Symbols are generally as in Figure 4.5, though here the open circles represents agricultural sites dominated by C4 crops which were not included in the linear regression in (b). The circled site in (b) is the Duke Old Field; when this site is neglected from the regression, $r^2 = 0.38$.

parameter in the grassland ecosystems. And again, some distinction is evident for the agricultural sites, which have significantly higher values of $\beta$ and $G_{Tref}$. However, these agricultural systems are short ecosystems, and the higher reference flux rates within this $PFT$ may be explained by the cross-site variation in $\beta$ and $G_{Tref}$ with $h$ (Figure 4.5), which has some theoretical justification as described in Section 4.5.3.

Thirdly, I asked "How much parameter variability can be attributed to inter-annual variation?" To answer this question, model inversions were performed a sub-
set of sites from Table 4.1 using an additional two years of data from each of these sites. As the distributions in Figure 4.4 illustrate, in some cases the parameters are reasonably well constrained from year to year; however, in many cases the posterior distributions for the same parameter barely overlap from year to year. Overall, the inter-annual variation is high, with parameter $C.V.s$ of 0.1 - 0.48 (mean 0.27). However, inter-annual variation is less than inter-PFT variation, and is considerably less for the reference rate parameters $\beta$ and $G_{Tref}$ (inter-annual $\langle C.V. \rangle = 0.19$ and 0.10, respectively, versus inter-PFT $\langle C.V. \rangle = 0.69$ and 0.27, respectively).

As the final step in this analysis, I asked ”Do relationships exist between model parameters and climatic variables or vegetation indices that are readily available at coarse scales? And could these relationships be useful in coarse-scale model parameterization schemes? Specifically, I tested for significant correlation between the model parameters and five indices: $h$, LAI, mean annual temperature, mean daytime growing season $D^{1/2}$, and mean annual precipitation. While many significant relationships existed, only a few had any explanatory power (i.e. $r^2 > 0.25$). These are the dependence of $G_{Tref}/LAI$, $\beta/LAI$, and $\gamma/LAI$ on canopy height ($r^2 = 0.49$, .44, and .38 when the wetland ecosystem is neglected) shown in Figure 4.5, and the relationship between $\lambda$ and long term $D$ for C3 ecosystems shown in Figure 4.7b ($r^2 = 0.27$ or 0.38 when site 101 is neglected from the analysis). In the remainder of this discussion section, following some remarks about the Gibbs sampling procedure and sources of error, I draw from existing theoretical frameworks in an effort to explain these relationships. Further, in Section 4.5.6, I use the results to quantify the accuracy of a purely PFT-based parameterization scheme to one that incorporates these inter-specific relationships. I conclude by making recommendations for future improvements in model parameterization.
4.5.1 Assimilation of eddy covariance data with Bayesian inversion techniques

While the Metropolis-Hastings algorithm has been used in many fields for decades, it has only recently been used in ecological applications. As other workers have pointed out (Van Oijen et al., 2005), Bayesian inversion with the M-H algorithm and other forms of Gibbs sampling has distinct advantages over traditional optimization techniques. First, Bayesian techniques allow for the estimation of posterior parameter distributions as opposed to single point parameter estimates. Posterior distributions that are well constrained with narrow ranges and standard deviations give confidence that the mean posterior estimates are close to the “true” parameter values for these datasets. Here, most of the posterior distributions were well constrained, with narrow confidence intervals (Table 4.5) and coefficients of variation of 0.1 - 0.15 for most of the NEE model parameters and <0.10 for the conductance model parameters (see Figure 4.1). One exception is the curvature parameter \( \theta_f \), which was generally more poorly constrained with an average posterior C.V. of 0.46. In section 4.5.5, I discuss in more detail some reasons why this particular parameter was relatively poorly constrained.

A second advantage of Bayesian inversion techniques is the ability to directly assess correlation among the model parameters. Here, within-site correlations among the posterior distributions was low for most parameters (i.e. \( r^2 \) less than 0.2). Some exceptions are the quantum yield parameter \( \alpha \), which was noticeably correlated with \( \beta, \gamma, \) and \( \theta_f \) for most sites (\( r^2 = 0.29, 0.58, \) and 0.43, respectively). The maximum assimilation rate \( \beta \) was also generally correlated with \( \theta_f \) (\( r^2 = 0.63 \)). Among the parameters of the conductance model, \( G_{Tref} \) was generally correlated with the light response parameter \( c_1 \) (\( r^2 = 0.44 \)). Some interdependence among model parameters is expected and has been demonstrated in previous studies (i.e. Owen et al. (2007)). If the posterior distributions for each site were simultaneously strongly correlated
and poorly constrained, then concerns about equifinality in the model parameters should be raised. However, given that the ranges of the posterior distributions were fairly narrow for nearly all of the model parameters, these mild correlations are not particularly troublesome. However, the existence of parameter correlations suggest that parameters may not be independent, and should caution ecosystem modelers against drawing closely related parameters from disparate sources.

A third advantage of the Bayesian framework is the ability to directly incorporate information from previous studies through the specification of prior distributions. Here, a lack of information about the nature of the parameter distributions in the literature motivated me to select prior distributions that were uniform and non-informative, allowing the inversion results to be driven primarily by the data. The posterior distributions for these parameters are close to, if not always, Gaussian (Figure 4.1), which is similar to results from other studies using Bayesian inference to estimate ecological model parameters (Knorr and Kattge, 2005). Therefore, I recommend a Gaussian form for the prior distributions of the parameters in future inversion studies.

4.5.2 Sources of error: Random measurement error, model selection, and LAI

The sources of error in eddy covariance measurements are well documented and may be attributed to a variety of factors, including mismatches between flux footprints and ecosystem dimensions (Detto et al., 2006; Schmid et al., 2000), mismatches between footprints for the net radiometer and the eddy covariance system (Baldocchi, 2003), instrument biases (Wilson et al., 2002), the separation of the gas analyzers and sonic anemometers, which acts as a low-pass filter (Oren et al., 2006), and, at some sites, micro-scale advective and drainage flows (Lee and Hu, 2002). Bias in eddy covariance data records have also been reported, though they appear to be more significant at night (Wilson et al., 2002) and should be less prominent in
these results, which rely on daytime measurements. Recently, random error measurement in eddy covariance measurements has been explicitly quantified using the paired-measurement technique (Richardson et al., 2006) employed here. The average estimates of random measurement error in $\sigma_\text{NEE} = 3.4 \, \mu\text{mol CO}_2 \, \text{m}^{-2} \, \text{s}^{-1}$ across all sites) and $\sigma_\text{LE} = 41 \, \text{W m}^{-2}$ are comparable to those reported by Richardson et al. (2006) for daytime conditions, suggesting that the quality control procedures employed here do not inject any additional uncertainty to the datasets beyond that which has been observed in other studies. Through the Bayesian inversions, error terms representing the mismatch between the data and the model was determined for $\text{NEE}$ and $\text{GT}$ (i.e. $\varepsilon_\text{NEE}$ and $\varepsilon_\text{GT}$). Because these error estimates incorporate both random measurement error (i.e. $\sigma(\delta)$) and error associated with model form and parameterization, comparing the $\varepsilon$ terms to the $\sigma(\delta)$ terms allows us to directly quantify the additional uncertainty attributed to the models themselves.

Across sites, $\varepsilon_\text{NEE}$ was about 40% higher than the random measurement error for $\text{NEE}$ (i.e. $\sigma(\delta)_\text{NEE}$), and the average value of $\varepsilon_\text{GT}$ was comparable to the average value of $\sigma(\delta)_\text{GT}$ (Table 4.3). Hence, while the non-rectangular model of Eq. 4.2 injects some additional uncertainty into the $\text{NEE}$ predictions, the model-data mismatch errors $\varepsilon_\text{NEE}$ are nonetheless close to the range of random measurement error alone as reported in Richardson et al. (2006). I also conclude that the Jarvis-model of Eq. 4.4 is a good model for eddy covariance derived conductance data, which indirectly confirm the assumptions that the majority of the latent heat flux measurements collected under high $\text{LAI}$ and more than two days after a rain even are dominated by canopy transpiration with minimal contamination from soil and interception evaporation.

Another source of error is uncertainty in $\text{LAI}$ measurements. Some of this analysis concerns relationships between model parameters and $\text{LAI}$ (i.e. Figure 4.6) or relies on model parameters normalized by $\text{LAI}$ (i.e. Figure 4.5). $\text{LAI}$ measurements
were only sporadically collected at many sites. Further, the error associated with LAI measurements, and in particular the remotely sensed LAI used to generate continuous LAI functions for many of the sites considered here, is known to be high (Breda 2003; Heinsch et al. 2006, Owen et al., 2007). Therefore, I expect that LAI dependent relationships will be noisier than if more accurate LAI estimation techniques and more frequent LAI estimates were available at the scale of the eddy-covariance footprint. Further, the attempt to normalize model parameters by LAI assumes that all leaf area present fully participates in gas exchange at all times, which is not likely to be the case in dense canopies with significant light interception and distributions. Leaf area density profiles could be used to instruct light interception models, which could then give a more accurate estimate of the effective sunlit leaf area in these canopies and could, presumably, decrease some of the variation in the cross-site relationships shown in Figures 4.5. Unfortunately, leaf area profiles are not currently available for most eddy covariance sites.

4.5.3 Inter-specific relationships between $G_{Tref}$, $\beta$, $\gamma$ and structural variables

$G_{Tref}$, $\beta$, $\gamma$ were all significantly related to canopy height when normalized by LAI, with $r^2 = 0.46$, 0.39, and 0.34, respectively. Removing the wetland ecosystem, where assimilation and respiration may be limited by standing water, improves correlation for $\beta/LAI$ to $r^2 = 0.46$. And limiting the analysis to non-managed, non-wetland ecosystems (i.e. forests) significantly improves correlation for the relationship between $\gamma/LAI$ and $h$ ($r^2 = 0.73$). This improvement may reflect the fact that the frequent removal of biomass from managed ecosystems artificially lowers the amount of organic substrate for autotrophic respiration.

A theoretical explanation for these trends follows from cohesion tension theory, as described in Chapter 3, Section 2. Specifically, noting that $G_{Tref}/LAI = G_{Sref}$, Equation 3.4 becomes
\[ G_{\text{Trf}/ \text{LAI}} \approx k_s \frac{A_s}{A_L} \frac{1}{h} (\Psi_{\text{leaf}} - \rho_w g h) , \tag{4.15} \]

which suggests that \( G_{\text{Trf}/ \text{LAI}} \) is, to an extent, a function of canopy height. To link \( \beta \) to \( h \), I rely on the Fickian-diffusion model for assimilation given in Eq 4.6. For reference environmental conditions,

\[ A_{\text{ref}} = \frac{G_{\text{Trf}}}{1.6} \cdot c_a \cdot \left(1 - \frac{c_i}{c_a}\right). \tag{4.16} \]

Noting that the when data are restricted to the reference environmental state (i.e. high light, non-limiting soil moisture, \( D \sim 1 \text{ kPa} \)), the assimilation rate approaches a maximum (i.e. \( A_{\text{ref}} \approx \beta \)). Hence, the arguments for an inter-specific height dependency for \( G_{\text{ref}/ \text{LAI}} \) may be extended to \( \beta/ \text{LAI} \). The linkage between \( \gamma/\text{LAI} \) and \( h \) is not as clear as this parameter incorporates both autotrophic and heterotrophic respiration. However, some similarity between \( \beta/ \text{LAI} \) and \( \gamma/\text{LAI} \) is expected given that leaf respiration is often assumed to be a constant fraction of assimilation rate (Collatz et al., 1991), and heterotrophic respiration has been recently shown to vary with \( A \) (Stoy et al., 2007).

In this analysis, the inverse of canopy height (1/\( h \)) explained approximately the same amount of variation in \( G_{\text{Trf}/ \text{LAI}} \) as \( h \) directly (\( r^2 = 0.44 \) vs 0.46, respectively), and 1/\( h \) explained less of the variation in \( \beta/ \text{LAI} \) and \( \gamma/\text{LAI} \) than \( h \) itself. This result is contrary to expectation and is in part driven by very low \( G_{\text{Trf}} \) estimates in some forests that are not necessarily the tallest canopies (Table 4.2, Figure 4.5). For some species, such as the spruce trees that dominate the Howland flux site (which has the lowest estimated \( G_{\text{Trf}/ \text{LAI}} \) in this study), \( A_s/A_L \) is known to decrease with canopy height (see previous chapter) and exceptionally low \( G_{\text{Trf}} \) in tall canopies is expected. However, \( A_s/A_L \) increases with \( h \) for many other species, compensating
for reductions in $G_{Tref}$ theoretically predicted from increasing $h$.

Surprisingly, the "ecosystem-scale" parameters $G_{Tref}$, $\beta$, and $\gamma$ were only weakly dependent on $LAI$ (Figure 4.7) despite the fact that all three parameters should depend directly on the total amount of leaf area present. Once again, self-shading, which limits the amount of leaf area in a canopy that actively participates in gas exchange at any one time, may confound directly relationships between these parameters and $LAI$. Further, the negative correlation between $LAI$ and $h$, which is particularly noticeable for forest ecosystems ($r^2 = 0.41$) may further weaken direct dependencies of these parameters on $LAI$ given the height dependencies revealed in Figure 4.5. I also note that this result is consistent with results from a previous study in which forest thinning from LAI = 8 to 6 m$^2$ m$^{-2}$ had no noticeable effect on CO$_2$ or water vapor flux (Vesala et al., 2005).

4.5.4 The sensitivity of stomatal conductance and water use efficiency to $D$

While the stomatal sensitivity parameter $b_d$ varies significantly across all sites, the mean value of this parameter is fairly well conserved across PFTs, varying by ca. 15% ($\langle b_d \rangle = 2.9, 3.3, 3.4, \text{and } 3.0$ for grasslands, agricultural sites, deciduous forests, and evergreen forests, respectively). I elected to model the sensitivity of $G_T$ to $D$ as a function of $D^{-1/2}$ because such a form has been recently justified in a new optimization theory framework describing vegetative gas exchange (see Katul et al. 2009). However, a Taylor series expansion of $D^{-1/2}$:

$$D^{-1/2} = 1 - 1/2 \log(D) + \left( -\frac{1}{2\log(D)} \right)^2 + \left( -\frac{1}{2\log(D)} \right)^3 + \ldots + \left( -\frac{1}{2\log(D)} \right)^n$$

(4.17)

directly links this function to the more widely used function $f_1(D) = 1 - m \log(D)$ (see Katul et al. 2009 for details). The sensitivity parameter $m$ of this function has also been shown to be fairly well conserved across species to values ranging from
0.5 - 0.6 (Oren et al., 1999b), and the higher order terms in the expansion of Eq. 4.17 become important only for very high $D$ which are not commonly encountered in these datasets.

As demonstrated in Section 4.2.3, $\lambda_L$ should also vary predictably with $D^{-1/2}$ according to Eq 4.11, at least on short time scales. For many sites, strong relationships emerged between these two variables from the hourly data ($r^2 > 0.60$), with an average hourly correlation of $\langle r^2 \rangle = 0.35$. Across C3 all sites, $D$ also explained a significant amount of the variation marginal water use efficiency (Figure 4.7b), implying the plants adapt to become more efficient in water use in more dessicating atmospheres. C4 ecosystems were neglected from this cross-site comparison because the concentration of CO$_2$ by bundle sheath cells near the photosynthetic apparatus of C4 leaves promotes the higher water use efficiency observed for these sites (see Figure 4.7b). Across all C3 sites, $\lambda$ averaged 958 $\mu$mol mol$^{-1}$, which is comparable to those reported in a previous study (Lloyd and Farquhar, 1994).

### 4.5.5 Issues in up-scaling from leaf to canopy

Both the non-rectangular hyperbola model for NEE and the Jarvis model for canopy conductance are implemented at the ecosystem scale allowing the parameters to be interpreted as ecosystem-scale parameters. However, linking these ecosystem-scale parameters to leaf-level physiological theory is instructional in the analysis and interpretation of the results. For example, efforts to link $G_{sref}$, $\beta$, and $\gamma$ to $h$ via cohesion-tension theory required the normalization of these parameters by $LAI$ to derive "leaf-level" analogues of these variables. As previously noted, simply dividing by $LAI$ ignores the impact of self-shading and light distribution in the canopy, and is likely to bias results for dense canopies, and in particular dense short canopies, where large portions of the leaf area are not sunlit and therefore do not participate effectively in gas exchange. While developing the light-interception models necessary
for a more robust determination of effective leaf-level parameters was beyond the scope of this study, these results suggest that leaf area profile measurements would be a useful ancillary data set at flux sites to better inform future modeling efforts.

Efforts to explain the observed relationships between $G_T$ and $\lambda$ to $D$ relied on matching parameters from Equation 4.2, which assumes that carbon assimilation is a hyperbolic function of $PAR$, to theory derived in part from the leaf-scale Farquhar
assimilation model Farquhar et al. (1980), which assumes that $A$ is linearly dependent on $PAR$ until it becomes Rubisco limited. Here again, patterns of vertical leaf area distribution and light penetration within the canopy are important emergent processes, and in this case can be used to directly link these two models. Specifically, Figure 4.8 shows the relationship between $A$ and $PAR$ predicted by the Farquhar model for a single leaf, formulated and parameterized after Novick et al. (2004). Also shown is the canopy-averaged value of $A$ when the Farquhar model is implemented in a multi-layered canopy, with light attenuation modeled by Beer’s law assuming a simple leaf area profile shown in the inset to Figure 4.8. The second curve clearly has a hyperbolic form which arises as the switch between light- and Rubisco-limited photosynthesis occurs at different light levels for each canopy sublayer. Thus, some theoretical justification for the hyperbolic form of Eq. 4.2 can come directly from the Farquhar model, which is the most widely used leaf-level model for vegetative CO$_2$ exchange.

The application of the Farquhar model in a multilayered canopy also helps explain why the curvature parameter $\theta_f$ and the maximum assimilation rate $\beta$ are among the most variable parameters in this analysis (Figure 4.3, Table 4.4 & 4.5). As temperature changes, so does the maximum assimilation rate, as both the maximum rate of Rubisco carboxylation and the potential rate of electron transport have temperature dependencies (Figure 4.8b). Thus, even within the relatively narrow temperature ranges associated with these study periods, some variability in $\beta$ should be expected with temperature which may explain why $\beta$ is associated with the most cross-site variation (Table 4.5). Furthermore, varying air temperature also affects the location of the transition point between light- and Rubisco- limited photosynthesis (Figure 4.8b), which directly impacts the curvature of the hyperbolic light response and may explain why the posterior distributions for $\theta_f$ are among the most poorly constrained posteriors (Table 4.4, Figure 4.1).
4.5.6  *PFT based parameterization – an evaluation and conclusions*

DGVMs are applied to large portions, if not all, of the global land surface, yet key physiological processes have been observed over only a small fraction of the earth’s terrestrial ecosystems. Hence, extrapolations from these measured sites to unobserved ecosystems, achieved through model parameterization, are a necessary part of quantifying and prognosing global carbon and water cycles. Historically, this parameterization has been accomplished on the basis of *PFT*, regardless of factors that vary significantly within a *PFT*, such as LAI, stand age and $h$, and climatic and edaphic conditions.

In this study, I show that the parameters of simple models for NEE and $G_T$ vary widely across sites, and that categorizing ecosystems by *PFT* does little to reduce this variation (Figure 4.3, Table 4.5). I then show that some of these parameters may be explained by climatic or structural indices that are readily available at coarse spatial scales; for example, the relationships between $G_{Tref}$, $\beta$, and $\gamma$ and canopy height (Figure 4.5), and the cross-site relationship between $\lambda$ and $D_{AVG}$ (Figure 4.7).

As a final step to this analysis, I tested the accuracy of a purely *PFT*-based parameterization scheme for these simple NEE and $G_T$ models against one that incorporates some of these inter-specific relationships. The average value of the model parameters were calculated for grasslands, agricultural sites, deciduous forests, and evergreen forests using half of the available sites (Group A). In the case of $G_{Tref}$, $\beta$, and $\gamma$, the parameter values were normalized by LAI. NEE and $G_T$ were then modeled at the remaining sites (Group B) using the *PFT*-averaged model parameters (*PFT-based* approach), with the normalized values of $G_{Tref}$, $\beta$, and $\gamma$ modified by the site-specific LAI. The mean relative error between the measured and modeled hourly flux rates was then calculated. Next, the process was repeated, but in this case $G_{sref}/LAI$ and $\beta/LAI$ were determined from linear relationships between these
Figure 4.9: The relative error between measured canopy conductance ($G_{Tref}$) and conductance modeled using the PFT-based approach (a) and the $h$-based approach (b). (c) and (d) show the relative error in measured NEE and modeled NEE using the PFT- and $h$-based approach, respectively. In both cases, the $h$-based approach takes advantage of the linear relationships between model parameters and $h$ shown in Figure 4.5. The mean values of the relative error ($\mu$) are shown for each model parameterization scheme.

parameters and $h$ determined from the Group A sites, and tested with the Group B sites by again modifying with site-specific LAI ($h$-based approach). In the case of forests (i.e. the non-managed, non-wetland ecosystems in this study), the parameter $\gamma/LAI$ was also determined from a derived linear relationship $h$ developed from the forested sites in Group A.

For the conductance simulations, the $h$-based approach reduced the average rel-
ative error across all sites from 0.31 with the *PFT-based* approach to 0.26, a change of -17% (Figure 4.9). For the *NEE* simulations, the magnitude of the average error across all sites was slightly higher with the *h*-based approach than the *PFT-based* approach (-0.08 vs 0.05, respectively); however, the *h*-based approach successfully decreased some of the largest relative errors and noticeably narrowed the distributions of errors for both the *GT* and *NEE* simulations.

Based on the results here and the large degree of uncertainty that still remains in the application of terrestrial ecosystem models, I recommend a re-evaluation of the *PFT*-based approach to parameterization. Physiological theory is increasingly being used to predict inter-specific functional convergence in plant functioning (Hickler et al., 2006; Meinzer, 2003; Oren et al., 1999b), and was used here successfully to explain cross-site relationships between certain model parameters and climatic and structural indices (Figures 4.5 and 4.7). Further, these relationships improved model performance over a *PFT*-based parameterization scheme (Figure 4.9). Future work should focus on identifying and characterizing similar inter-specific relationships that may reduce uncertainty associated with *DGVM* parameterization schemes. I note that more frequent ground-based *LAI* measurements and leaf area profile measurements could facilitate such an effort, as the vertical attenuation of light within the canopies has important consequences for linking ecosystems scale parameters to leaf-level theory. More accurate remotely-sensed *LAI* measurements and schemes to estimate leaf area profiles over coarse spatial scales would also improve coarse-scale model outputs. More complete records of soil moisture at FLUXNET sites are also necessary to properly account for the impact of moisture deficits on model parameters, which is a necessary step to permit the application of the results from this analysis over large land areas. I also note that other work has recently focused on variation in parameters with time (Chen et al., 2008), and these results shows that they may vary considerably from year to year (Figure 4.4, Table 4.5). Thus, care
must be taken in employing parameterization schemes that do not properly account for inter-annual variability, and sources of this inter-annual variation should continue to be an active research topic.
5

Elevated atmospheric CO$_2$ enhances resin production in field-grown pine trees

5.1 Introduction

Bark beetles (Curculionidae: Scolytinae) and their fungal associates are among the most prominent natural disturbance agents in coniferous forests and can have profound economic and ecological impacts (Kurz et al., 2008; Raffa et al., 2008). Recently, large scale bark beetle outbreaks have caused extensive tree mortality in several independent infestations, including the ongoing Mountain Pine Beetle outbreak in British Columbia, Canada (Kurz et al., 2008), the recent outbreak of Pinyon ips. beetles in southwestern North America (Raffa et al., 2008), and the large outbreak of spruce beetles in Alaska in the 1990’s (Berg et al., 2006). These outbreaks have focused much political and scientific attention on the interactions between bark beetles and their host species (Negron et al., 2008) and the potential for host-insect relationships to change in future climates.

Beetle-forest interactions are sensitive to air temperature, and the range of many beetle species appears to be limited by low temperatures in winter (Williams and...
Some studies suggest an increase in beetle activity or an expansion in their geographic range in warmer climates or cite higher air temperatures as a co-factor in recent beetle outbreaks (Berg et al., 2006; Kurz et al., 2008; Logan et al., 2003; Raffa et al., 2008). Beetle outbreaks may be further promoted by the increasing distribution of managed, mono-cultural pine forests, which increase the density of suitable host trees while decreasing the abundance of natural beetle predators (Fettig et al., 2007; Kurz et al., 2008).

Ongoing climate and land use change set the stage for more frequent and perhaps more severe bark beetle outbreaks in coming decades (Raffa et al., 2008; Gan, 2004). However, future forest-pest dynamics cannot be divorced from host defense mechanisms, which have not been well studied for pine systems growing in conditions representative of future climates. Resin, which flows through a series of ducts formed in the phloem and ejects or entombs attacking beetles (Trapp and Croteau, 2001), is the primary defense against bark beetle attack in pine species.

Early work by Waring and Pitman (1985) and Christiansen et al. (1987) suggested that resin production is linked to tree vigor (or the annual stem biomass increment per unit leaf area), and that any environmental stress may lead to a scarcity of photosynthates and a decrease in tree resistance to beetle attack. More recent work conducted within the framework of the "growth-differentiation balance" (GDB) hypothesis for plant defense (Herms and Mattson, 1992) refined these earlier conclusions by suggesting that any environmental stress that limits photosynthesis relative to growth processes may lead to a decrease in resin production. However, according to the GDB framework, environmental stressors that limit growth more than photosynthesis (for example, mild nutrient or moisture deficits) may actually increase resin production (Lorio et al., 1995; Knebel et al., 2008; Stamp, 2004; Lombardero et al., 2000a) by increasing the total pool of available photosynthates. In other words, resin production should be related to the balance between total carbon
assimilation (A) and the amount of assimilated photosynthates allocated to growth processes (G), noting that each may respond differently to various environmental drivers.

Here, the impact of elevated CO\textsubscript{2} and other factors including tree size, time of year, and fertilization on constitutive resin flow is explored in the Duke Free Air Carbon Enrichment (FACE) facility located in a loblolly pine stand near Durham, NC (USA). This study is the first to report on the impact of eCO\textsubscript{2} on resin flow in mature field grown pine trees. From GDB theory, I expect that resin flow and biomass growth should be inversely correlated on short time scales provided A is constant or at least does not vary as much as G. Over longer time scales, I expect resin flow to be higher in dominant versus suppressed trees as the former have more ready access to sunlight, which considerably increases A. I also expect greater resin flow in elevated versus suppressed trees, as elevated CO\textsubscript{2} enhances A more than G in the Duke FACE facility owing to growth restrictions imposed by nutrient deficits (Oren et al., 2001a). Fertilization affects both assimilation and growth, and the impact of fertilization on resin production will depend on whether nutrient amendments enhance A more than G, or vice versa.

5.2 Materials and methods

5.2.1 The Duke FACE experiment

The Duke FACE facility is located in a Pinus taeda L. (loblolly pine) plantation in the Duke Forest, North Carolina, U.S.A. (35° 48’ 51” N, 79° 05’ 39” W). The plantation was established in 1983 following a clear cut and a burn. Seedlings were planted at 2.0 m by 2.4 m spacing and ecosystem development has not been managed after planting, resulting in the development of a diverse understory comprised of Liquidambar styraciflua L. and over 40 other deciduous broadleaf species. The soil is classified as moderately low-fertility acidic clay-loam (Enon Series). The mean
annual temperature is 15.5 °C and mean annual precipitation is 1,145 mm.

The FACE experiment includes eight circular plots 30 m in diameter – four that experience ambient CO$_2$ (aCO$_2$) conditions and four that experience elevated CO$_2$ (ambient +200 ppm, eCO$_2$) conditions (Hendrey et al., 1999). Two of these plots (one ambient plot and one elevated plot) comprise the FACE prototype complex, which was established in 1993. The remaining FACE plots were established in 1996 as blocked replicates. In 1998, the prototype FACE ring and its reference plot were split and one half of each received yearly N fertilization (11.2 g N m$^{-2}$ y$^{-2}$, Oren et al. 2001a). Identical fertilization treatments began in the replicated FACE plots in 2005. This split-plot amendment created 16 subplots within the FACE facility: 4 aCO$_2$, unfertilized (AU) plots, 4 aCO$_2$, fertilized plots (AF), and four each of unfertilized and fertilized eCO$_2$ plots (EU and EF, respectively). In this study, the prototype and replicated FACE plots are treated as one replicated experiment following initial analysis that failed to show a significant block effect on resin flow (P = 0.62 from a split-plot ANOVA with block as a random effect) and prior results that strongly reject the hypothesis that the prototype and replicated FACE complexes are not from the same study (McCarthy et al., 2007).

5.2.2 Resin measurements

Resin mass flow ($m_R$) was measured monthly in 5 trees (3 canopy dominants and 2 canopy subdominants) within each of the 16 treatment sub-plots with two measurements per tree per month. Resin was measured by removing an 8 mm disk of bark to the phloem-xylem interface at breast height on opposite sides of the tree, and installing trays below the wound to direct resin into pre-weighed collection vials for a 24-hour period. Each month, a new piece of bark was removed at a location that was approximately 5 cm away from previous sampling locations. Following standard practice (Klepzig et al., 2005; Knebel et al., 2008), samples were assumed to be
independent with previous sampling times having negligible impact on subsequent resin collection (Knebel et al., 2008). All collection vials were installed between the hours of 0800 - 1200 to minimize any time-of-day effects on resin flow. Results are expressed as gram of resin per day normalized by the area of the wound (0.50 cm²).

5.2.3 Ancillary data

A suite of meteorological measurements were also collected during the study period. Half-hourly precipitation was measured using a tipping bucket (TI, Texas Instruments, Austin, TX, USA) and $T_a$ was measured with an HMP35C Ta/RH (Vaisala, Finland) positioned at 2/3 canopy height. Gross canopy assimilation was measured using an eddy-covariance system comprised of triaxial sonic anemometers (CSAT3, Campbell Scientific, Logan, UT, USA) and open-path infrared gas analyzers (IRGA, LI-7500, Li-Cor, Lincoln, NE, USA) following measurement, quality control, and gap-filling protocols described in Chapter 2. Diameter at breast height was measured bi-weekly during the growing season and monthly during the rest of the year with dendrometer bands. Aboveground net primary productivity of wood biomass ($ANPP_W$) was estimated from dendrometer measurements and laser-based tree height measurements after McCarthy et al. (2006a).

5.2.4 Statistical considerations

The effects of eCO$_2$ and fertilization were assessed via a split-plot analysis of variance (ANOVA), where CO$_2$ concentration and fertilization status were the main and split-plot effects, respectively, and blocks were used as replicates. Measurements made on multiple months were analyzed by repeated measure ANOVA. A nested ANOVA design was initially employed to assess the impact of size class and block on $m_R$. The ANOVA analyses were performed using SAS (version 9.1, Cary, NC, USA), and results were compared to the output from a Bayesian hierarchical inversion.
performed in Matlab (Mathworks Inc., Boston, MA, USA), which was also designed to test the effects of CO$_2$ concentration, fertilization, and sampling date on resin flow (see Appendix B). Significance levels for mean responses between different treatment groups were assessed using a one-tailed t-test for unknown means with equal variances and a null hypothesis of no difference between the means.

5.3 Results and Discussion

Resin flow (shown in Figure 5.1a as monthly averages across all treatments) is highest during the growing season (April - September), with a local minimum observed in May and a maximum observed in August. $ANPP_W$, conversely, reaches a maximum in May and a local minimum in August (Figure 5.1b). Gross canopy assimilation, estimated in one of the ambient plots using the eddy technique, was high and relatively constant throughout the peak of the growing season (Figure 5.1c). Mean air temperature and precipitation during the study period were found to be near their long-term means, though a minor soil moisture deficit developed at the site in August and September.

For the duration of the study period, resin flow was significantly higher in dominants versus suppressed trees when data were marginalized across CO$_2$ and fertilization treatments ($P < 0.001$ from an ANOVA with size class nested within subplots), confirming the hypothesis from GDB theory. To separate size class from other treatment effects, the remainder of the analysis is performed separately for dominant and suppressed individuals.

The data in Figure 5.1a and 5.1b suggest an inverse relationship relationship between $m_R$ and $ANPP_W$ during the growing season as predicted from GDB theory. This negative correlation was in fact observed within all treatment classes (Figure 5.2a) with particularly strong negative relationships observed for dominants ($r^2 = 0.64$ to 0.90 when data are averaged within each treatment). These relationships
do not appear to be driven by assimilation rates, as $A$ (at least in ambient plots) was high throughout the peak of the growing season (Figure 5.1c) even as $ANPP_W$ declined significantly (Figure 5.1b).

Moreover, the slope parameter representing the decline in $m_R$ with $ANPP_W$ ($\partial m_R/\partial ANPP_W$) was closely related to the intercept parameter ($m_{R,max}$). The relationship between $\partial m_R/\partial ANPP_W$ and $m_{R,max}$ is different for suppressed versus dominant trees, which can be explained by greater $A$ in dominant trees allowing for
Figure 5.2: Over short time scales, resin flow \( (m_R) \) is inversely related to the net primary production of aboveground biomass \( (ANPP_W) \) within all treatment sub-classes (a). The lines show the linear relationship between these two variables derived from the mean monthly values for all trees within each treatment sub-class. The correlation coefficient associated with the mean monthly data used to generate the linear relationships are shown in the legend. The relationship between the intercept parameter \( m_{R,max} \) and the slope parameter \( (\partial m_R / \partial ANPP_W) \) for each line is shown in (b).

relatively more resin production even when \( ANPP_W \) is high. These results suggest a convergence in the relationship between \( m_R \) and growth processes across CO\(_2\) and nutrient treatment regimes over short time scales. Hence, eCO\(_2\) and fertilization primarily impact \( m_R \) by determining the long-term reference flow rates (i.e. \( m_{R,max} \)).

According to the results of a repeated measures ANOVA, eCO\(_2\) significantly enhanced resin flow in canopy dominants with no significant fertilization impact (Table
5.1). However, as the results in Figure 5.2 illustrate, the eCO₂ enhancement was largely limited to unfertilized trees. The mean value of $m_{R,max}$, determined from estimates of $m_{R,max}$ within each subplot, increased from $2.4 \pm 1.5$ g cm$^{-2}$ d$^{-1}$ for unfertilized dominants growing in aCO$_2$ to $5.8 \pm 2.2$ for unfertilized trees in eCO$_2$ – an increase of >100%. Among fertilized trees, the mean value of $m_{R,max}$ was similar in both aCO$_2$ (3.9 ± 0.9 g cm$^{-2}$ d$^{-1}$) and eCO$_2$ (4.2 ± 1.9 g cm$^{-2}$ d$^{-1}$) and the difference is not significant (P = 0.61). Previous work at the Duke FACE experiment has shown that eCO$_2$ stimulates photosynthesis by 40-60% in unfertilized plots even after 9 years of treatment (Crous et al., 2008; Maier et al., 2008). Net primary productivity is also enhanced by eCO$_2$, but only by ca. 18 - 25 % (McCarthy et al., 2006a) due to constraints on growth imposed by nutrient limitation (Oren et al., 2001a). Hence, unfertilized trees grown in elevated plots have excess unallocated carbon, and the results from this study suggest greater allocation to resin production in these trees, consistent with GDB predictions.

The impact of nutrient limitation on $m_R$ is also evident among plots experiencing similar CO$_2$. For unfertilized plots, mean $m_R$ declined with increasing N availability (Figure 5.3a) in both aCO$_2$ and eCO$_2$ and resin flow was highest in plots with the most severe growth restrictions. In contrast, no significant relationship between $m_R$ and N availability was observed among fertilized plots (Figure 5.3b).

The GDB framework provides some explanation for the variation in $m_{R,max}$ between fertilized and unfertilized trees grown in both aCO$_2$ and eCO$_2$. Fertilization simultaneously stimulates photosynthesis and growth processes, though the degree of enhancement of each may vary across stands. If the nitrogen treatment enhanced growth in elevated plots more than it enhanced photosynthesis (which is already stimulated by eCO$_2$), then $A - G$ would be lower in fertilized versus unfertilized eCO$_2$ plots, explaining the observed though insignificant decreases in $m_{R,max}$ (2.1 ± 1.0 versus 2.9 ± 1.1 for fertilized versus unfertilized eCO$_2$ trees, respectively, P =
### Table 5.1: The probability values for the impacts of the two atmospheric CO$_2$ concentration levels (C), the two fertilization levels (N), and the eight measurement dates (M) on resin production ($m_R$) from a repeated measures, split-plot ANOVA with CO$_2$ concentration as the main effect and fertilization treatment as the split effect. The ANOVA was run independently for dominant and suppressed trees, where dominant trees are those that have at least 50% of leaf area exposed to direct sunlight, and suppressed trees are completely or nearly overtopped by dominants. Significant effects are highlighted.

<table>
<thead>
<tr>
<th>Effect</th>
<th>df</th>
<th>$m_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dominant (n = 60)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0.03</td>
</tr>
<tr>
<td>N</td>
<td>1</td>
<td>0.78</td>
</tr>
<tr>
<td>C x N</td>
<td>1</td>
<td>0.13</td>
</tr>
<tr>
<td>M</td>
<td>7</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>C x M</td>
<td>7</td>
<td>0.48</td>
</tr>
<tr>
<td>F x M</td>
<td>7</td>
<td>0.47</td>
</tr>
<tr>
<td>C x F x M</td>
<td>7</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Suppressed (n = 40)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0.97</td>
</tr>
<tr>
<td>N</td>
<td>1</td>
<td>0.60</td>
</tr>
<tr>
<td>C x N</td>
<td>1</td>
<td>0.63</td>
</tr>
<tr>
<td>M</td>
<td>7</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>C x M</td>
<td>7</td>
<td>0.45</td>
</tr>
<tr>
<td>F x M</td>
<td>7</td>
<td>0.45</td>
</tr>
<tr>
<td>C x F x M</td>
<td>7</td>
<td>0.02</td>
</tr>
</tbody>
</table>

In ambient plots, if fertilization enhanced $A$ more than $G$, then an increase in $A - G$ in fertilized versus unfertilized aCO$_2$ plots could explain the observed increase in $m_{R,max}$ between these two treatments (1.93 ± 0.4 versus 1.2 ± 0.8 for fertilized versus unfertilized aCO$_2$ trees, respectively, $P = 0.07$).

These theoretical explanations for the impact of fertilization on $m_R$, at least in the case of eCO$_2$, are not entirely supported by the $ANPP_W$ data. Fertilization produced an insignificant and only marginal increase in $ANPP_W$ (i.e. $< 1$ g C tree$^{-1}$ day$^{-1}$) for the trees in this study. Of course, the increment in aboveground woody
Figure 5.3: Resin flow ($m_R$) decreases as a function of native N availability in dominant trees growing in unfertilized plots (a). The symbols are as in Figure 5.2 and show the mean value within each subplot for the entire study period, with the error bars showing the standard error for the mean estimated for all available measurements with each subplot. By contrast, no clear relationships are evident for trees growing in fertilized plots (b).

Biomass does not incorporate all growth processes, which also include increments in leaf area and root biomass. Further, the carbon requirement for $m_R$ is small and could be influenced by even slight changes in $A$ or $G$. Nonetheless, the variable and non-significant impact of fertilization treatments on $m_R$ observed in this study are fairly representative of conflicting results reported in the wider literature (Klepzig et al., 2005; Knebel et al., 2008) and highlight a need for additional studies investigating resin flow under a wide range of nutrient conditions.
5.4 Conclusions

The Duke FACE experiment allowed us to test, for the first time, the impact of elevated atmospheric CO_2 concentration on the production of an important defensive compound (i.e. resin) in field-grown pine trees. I found that elevated CO_2 significantly enhanced resin production in canopy dominant trees (Table 1.1, Figure 5.2) with most of the enhancement limited to unfertilized trees, for which eCO_2 enhanced \( m_R \) by \( >100\% \). I proposed that the large eCO_2 enhancement in unfertilized trees could be explained by eCO_2 enhancements to A concurrent with nutrient limitations to G, which together increase the pool of available photosynthates available for resin synthesis. I also showed that the range of N availability in unfertilized subplots explains inter-plot differences in resin flow rates (Figure 5.3). Over short time scales, I demonstrated that resin flow is inversely correlated with ANPP_W within all treatment classes (Figure 5.2a), and that the rate of change of \( m_R \) with ANPP_W (\( \partial m_R / \partial ANPP_W \)) is strongly related to the maximum flow rate (\( m_{R,max} \)) within tree size classes (Figure 5.2b).

These results are largely explained by the growth-differentiation balance hypothesis, which in turn is consistent with the early findings of Waring and Pitman (1985) if tree vigor is re-defined as the balance between total available photosynthates and those required for growth processes (i.e. A - G). Originally, Waring & Pittman defined vigor primarily as a function of G. I note, however, that certain processes that enhance A more than G (for example, eCO_2) or that decrease G more than A (for example, ontogentical control of G over the course of the growing season) would promote a direct relationship between (A - G) and G alone.

The observed eCO_2 enhancement in resin production may buffer future forests against projected increases in beetle populations associated with a warmer climate (Berg et al., 2006; Kurz et al., 2008; Raffa et al., 2008), at least in forests with
low- to mid-range fertility. Certainly, the response of resin production and flow to future climate conditions cannot be ignored in efforts to forecast the magnitude and frequency of future pine beetle outbreaks. The results here also show that resin production is dynamic over the course of the growing season, increasing from a mean flow rate across treatments of 1.5 g m$^{-2}$ d$^{-1}$ in May to 3.3 g m$^{-2}$ d$^{-1}$ in August. Hence, the impact of future climate conditions on the timing of bark beetle outbreaks is also important to consider in forecasts. Fortunately, over short time scales, the close coordination of $m_{R,\text{max}}$ and growth demonstrated here could facilitate efforts to represent the dynamics of resin flow in model applications.
Effects of water table depth variation on ecosystem-scale fluxes of greenhouse gases in a large restored temperate wetland

6.1 Introduction

Over the course of the past 200 years, 40% of wetland ecosystems in the Southeastern U.S. have been drained and converted to other land uses such as forestry and agriculture (Dahl, 1990). In more recent years, economic pressures have led to large scale agricultural abandonment in this region (Wear and Greis, 2002). In recognition of the numerous ecosystem services provided by wetlands, including flood control and nutrient removal (Zedler, 2003), a number of these former agricultural areas are being restored to wetland ecosystems through compensatory mitigation (Morse, 2010). However, some uncertainty remains regarding potential tradeoffs between the known water quality benefits of wetlands and the emissions of greenhouse gases (GHGs) such as CO₂, CH₄, and N₂O, which are prevalent in the anoxic soil environments that characterize these ecosystems (Dodla et al., 2009; Verhoeven et al., 2006).

Much of this uncertainty is related to the large variability in wetland GHG fluxes,
which can range across orders of magnitude (see Table 1.1) with both time and space even within study sites (Bubier et al., 2005; Conrad, 1996; Groffman et al., 2009; Johansson et al., 2003; Longoria-Ramirez et al., 2003; Whalen, 2005). The presence of small areas ("hot spots") of high GHG production in wetland soils are believed to be an important factor in the high variability of GHG exchange (Groffman et al., 2009), as is variation in water table depth and related variation in water chemistry and redox potential (Bubier et al., 2005; Dusek et al., 2009; Jauhiainen et al., 2005; Smemo and Yavitt, 2006).

The most frequently employed method for measuring these fluxes is soil chamber measurements, which are not well suited for highly heterogeneous ecosystems given the small measurement volume of the chambers. Further, chambers are somewhat invasive structures that may create micro-environmental disturbance leading to measurement bias (Lai, 2009). And in areas of tall vegetation, they can not appropriately capture important GHG transformations that occur in aboveground plants, such as photosynthetic carbon assimilation and the transfer of GHGs from the soil to the atmosphere through the vascular systems of aboveground plants (i.e. Le Mer and Roger 2001).

Two micro-meteorological methods are employed here in a large restored wetland to directly estimate fluxes of CO$_2$, H$_2$O, CH$_4$, and N$_2$O at spatial scales that are several orders of magnitude larger than typical chamber area (i.e. 10$^2$ to 100$^2$ meters), permitting a de facto integration over small-scale heterogeneities and direct observational inclusion of above-ground processes. These atmospheric measurements are not free from their own difficulties, assumptions, and technical issues, but may nonetheless provide some constraints on biosphere-atmosphere fluxes of GHGs.

These atmospheric measurement systems were deployed together in several hydrologically distinct areas of the wetland. The compass of this work is centered on using this array of flux estimates to address the following questions:
1. How does water table depth impact GHG fluxes?

2. Do rapid changes in water table depth impact GHG fluxes?

3. What is the uncertainty in the flux estimates when inferred from aerodynamic methods in a heterogeneous study site?

With regards to Questions 1 and 2, thermodynamic theory may be used to formulate some hypotheses about microbial activities. The biological activity of aerobic, nitrifying, denitrifying, and methane-producing soil microbes is largely controlled by the redox potential of the soil matrix (Schlesinger, 1997). In oxic soils with high redox potential, heterotrophic respiration of CO$_2$ is the favored metabolic pathway. Nitrifiers, which are aerobic chemoautotrophic bacteria, are also present in oxic soils and derive energy from the oxidation of NH$_4^+$ to NO$_2^-$ and NO$_2^-$ to NO$_3^-$, with some N$_2$O produced as an intermediate by-product (Paul and Clark, 1996). Nitrifiers grow slowly and tend to be poor competitors for NH$_4^+$ with respect to heterotrophic microbes (Megonigal et al., 1996).

In more anaerobic soils with lower redox potentials, facultative denitrifying microbes use NO$_3^-$ as a terminal electron acceptor in the absence of O$_2$, ultimately producing N$_2$ (Knowles, 1982), with some N$_2$O again produced as a by-product. Anaerobic conditions also favor methanogenesis, which includes two pathways of CH$_4$ production by Archaea. In aerobic areas of wetland ecosystems with high redox potentials, CH$_4$ may actually be consumed by methanotrophic bacteria, a process which may considerably mitigate CH$_4$ emissions from more reducing wetland soils (Conrad, 1996; Whalen, 2005).

Soil redox potential is largely related to water table depth. In flooded soils, redox potential is low and the energetically unfavorable metabolic processes are more likely. In non-flooded soils, the aerobic processes (i.e. heterotrophic respiration, nitrification,
and methanotrophy) are favored. Hence, it is anticipated that aerobic respiration will dominate in oxic soils without significant effluxes of $N_2O$ and $CH_4$ in these locations during dry periods. Effluxes of $CO_2$ at these sites may be high, particularly during periods of low photosynthetic assimilation. However, during daylight hours when photosynthetic $CO_2$ assimilation is high, these areas may function as sinks of $CO_2$. The dryland areas of the wetland may function as methane sinks if methanotrophic bacteria are present.

In flooded areas of the wetland, $CO_2$ production by aerobic respiration as well as $CO_2$ assimilation by vegetation should be limited. Significant effluxes of $CH_4$ and $N_2O$ are expected in flooded areas of the wetland where energetically unfavorable metabolic pathways are promoted by low redox potentials. With this background, I now proceed to discuss the measurement techniques for estimating the GHG fluxes.

### 6.2 Theoretical considerations

Two micro-meteorological techniques to estimate greenhouse gas fluxes were applied in a large restored wetland in Eastern North Carolina, USA, during the 2008 and 2009 growing season. The first is the eddy covariance methodology (hereafter referred to as $EC$), which is used to measure high-frequency fluxes of $CO_2$ and $H_2O$ as well as other energy fluxes at two locations in the study site from instrumentation mounted on a Mobile Meteorological Station ($MMS$, Figure 1). The $MMS$ was positioned in a dryland portion of the study site in 2008 and an inundated portion of the site in 2009. The other methodology, hereafter the Scalar Gradient method ($SG$), is a variant of the flux-gradient technique that relies on vertical gas concentration profiles collected from a tower situated in an upland portion of the wetland ($TWR_A$) and from a permanently inundated portion of the wetland ($TWR_B$, see Figure 1). In this section, the theoretical details of these techniques are discussed. Following standard meteorological convention, a negative flux denotes uptake of the scalar by
the ecosystem.

6.2.1 Eddy covariance (EC) approach: Overview of assumptions and limitations

The EC methodology is frequently used to determine biosphere-atmosphere exchanges of mass and energy and detailed overviews of the methodology are available elsewhere (Baldocchi, 2003). Briefly, the governing equation for the mean scalar continuity equation above a canopy of height \( h \) is given by (Stull, 1988):

\[
\frac{\partial \bar{s}}{\partial t} + \mathbf{u}_j \frac{\partial \bar{s}}{\partial x_j} = - \frac{\partial}{\partial x_j} \left( \mathbf{u}_j' \bar{s}' - v_c \frac{\partial \bar{s}}{\partial x_j} \right)
\]

where the subscript ' \( j \)' denotes the three principle directions (i.e. \( x_1, x_2, \) and \( x_3 \) or \( x, y, \) and \( z \)), \( s \) is the instantaneous concentration (mass volume) of a scalar quantity, \( \mathbf{u}_j \) is the instantaneous air velocity with directional components \( u, v, \) and \( w, v_c \) is the molecular diffusivity of scalar \( s \), and primed quantities (i.e. \( u_j', \ s' \)) denote fluctuations about the mean state. For stationary (\( \partial \bar{s}/\partial t = 0 \)) and planar-homogeneous (\( \partial(\cdot)/\partial x = \partial(\cdot)/\partial y = 0 \)) flow at high Reynolds numbers (\( |v_c \bar{s}/\partial x_j| \ll |u'_j \bar{s}'| \)), and in the absence of large-scale subsidence (\( w_3 = \bar{w} \approx 0 \)), the mean scalar budget equation reduces to \( \partial \bar{w}' \bar{s}'/\partial z \approx 0 \). This simplified equation may be integrated with respect to \( z \) up to the measurement height (\( z_{MEAS} > h \)), yielding

\[
\bar{w}' \bar{s}'(z_{MEAS}) = \bar{w}' \bar{s}'(h).
\] (6.1)

These assumptions are the basis for the ‘constant-flux’ layer, which permits turbulent flux measurements above the canopy to represent fluxes from the ecosystem. Moreover, if these same assumptions are applied to a large enough control volume inside the canopy, then \( \bar{w}' \bar{s}'(h) \) represents all the sources and sinks within this control volume, including all contributions from the canopy and the ground, provided the
dispersive fluxes\textsuperscript{1} are negligible compared to the turbulent fluxes.

In recent decades, high-frequency gas analyzers have been developed that permit the measurement of CO\textsubscript{2} and H\textsubscript{2}O concentration fluctuations at frequencies that match turbulent time scales (i.e. 5 - 20 Hz), and these instruments may be co-located with sonic anemometers to directly estimate the turbulent flux term $\overline{w's'}$. This term is representative of the surface flux for each scalar (i.e. $F_{CO2,EC} = \overline{w'[CO2]}$) with similar formulations for H\textsubscript{2}O and heat) provided the assumptions described above are met, and provided the footprint of the eddy covariance system lies within the ‘fetch’ of the study ecosystem. The flux "footprint" is an estimate of the planar extent of the land area that contributes to atmospherically derived flux estimates. In this study, the flux footprint for the EC instrumentation is calculated as the peak of the source-weight function ($x_p$) using an analytical model based on Lagrangian dispersion theory (Hsieh et al., 2000).

6.2.2 The Scalar Gradient method (SG)

Following Monin-Obukhov similarity theory above the canopy, the vertical flux of any scalar quantity ($F_S$) may be linked to the mean vertical concentration profile of that scalar by a first-order closure approximation scheme:

$$F_S = \overline{w's'} = -K_s(z)\frac{\partial \overline{s}}{\partial z} \quad (6.2)$$

where $K_s$ is the eddy diffusivity for scalar $s$ at height ($z$). From similarity theory, it is often assumed that the diffusivities for heat ($K_h$) and momentum are similar (i.e. the turbulent Schmidt number is unity), and the eddy diffusivity of heat is also

\textsuperscript{1} Dispersive flux terms are formed when the time-averaged mean continuity equation is spatially averaged within the canopy volume. These fluxes represent a contribution to scalar transfer arising from the spatial correlation of the time-averaged vertical velocity and scalar concentration within a horizontal plane embedded in the canopy. They have been shown to be small in a number of laboratory and field experiments (Poggi et al., 2004; Poggi and Katul, 2008).

148
identical to those of trace gases like CO$_2$ (Ohtaki, 1985; Laubach and Kelliher, 2004). Therefore, in this analysis, I assume that the eddy diffusivity is identical for heat and all of the scalar trace gases (i.e CO$_2$, CH$_4$, and N$_2$O) measured on the profile towers. A large number of theoretical and field studies have demonstrated that such equality may be questionable (Warhaft; Lang et al. 1983; Katul and Hsieh 1999 and others). However, in the absence of any alternative approach, equality among the diffusivities is assumed (i.e. $K_s = K_h = K_{CO_2} = K_{CH_4} = K_{N_2O}$).

With this assumption, the flux and vertical gradients of the GHGs may be directly linked to the sensible heat flux ($H$) and vertical gradients for mean air temperature. For example, the flux of CO$_2$ ($F_{CO_2,SG}$) may be expressed as:

$$F_{CO_2,SG} = \frac{H}{\rho c_p} \left( \frac{\partial T_a}{\partial z} \right)^{-1} \left( \frac{\partial [CO_2]}{\partial z} \right)$$  \hspace{1cm} (6.3)

with analogous formulations for $F_{CH_4,SG}$ and $F_{N_2O,SG}$. Here, $\rho$ is air density, $c_p$ is the specific heat capacity of dry air, $T_a$ is temperature, and $[CO_2]$ is CO$_2$ concentration (mg m$^{-3}$). To use equation 6.3 in practice, the gradients must be estimated from a finite number of samples. I assumed a log-log variation in gas concentration and temperature over the narrow range in $z$ that bounds the temperature and concentration profiles. For example, the relationship between CO$_2$ and $z$ is assumed to be:

$$\log [CO_2] = A_z \log (z - d) + B_z,$$  \hspace{1cm} (6.4)

where $d$ is the zero plane displacement. When differentiated with respect to $z$ and referenced to the measurement height, Eq. 6.4 gives

$$\frac{\partial [CO_2]}{\partial z} = [A_z B_z (z_{MEAS} - d)]^{B_z-1}$$  \hspace{1cm} (6.5)
6.3 Materials and methods

6.3.1 Site description

The study site is a former agricultural field within the Timberlake Restoration Project (TRP), owned by Great Dismal Swamp Mitigation Bank, LLC and located in the Albemarle-Pamlico Peninsula in Eastern North Carolina, U.S.A. (35° 54' 22" N, 76° 09' 25' 'W). Historically, the TRP land was a swamp forest that served as the headwaters for coastal blackwater streams that flow into the Alligator River. The land was cleared and drained in the 1970’s and operated as a corn and soybean agricultural system until 2004, when restoration of the site to a riverine wetland was initiated. Ditches were filled, canals were plugged, and 750,000 live samplings from eight species of obligate and facultative wetland trees were planted across the site. Disabling the downstream gate-pump system in February 2007 completed the final step in the restoration process, reinstating the wind tide driven hydrologic regime.

Soils on the site include Hyde loam, Ponzer muck, Roper muck, Belhaven muck, Weeksville silt loam, and Pungo mucks: all of which are classified as rarely flooded, very poorly drained hydric soils. However, deforestation, drainage, and tillage associated with agriculture have caused much of the original peat and muck to be oxidized over time, such that the soils on the property may now be borderline mineral soils. The elevation in the TRP ranges from -0.4 m to 5.1 m above sea level.

The vertical gas concentration profiling towers (i.e. TWR_A and TWR_B) were located in two vegetatively and hydrologically distinct areas of the study site (Figure 6.1). The area around TWR_A is populated with a variety of native grasses and shrubs reaching a height of approximately 1 m. In 2008, the water table near TWR_A was well below the surface, and soils within the upper 15 cm of the soil profile were unsaturated. In 2009, the water table depth was several inches above the surface of most of the ground area, though some ground areas emerged above the water surface.
due to micro-topographical variation.

The area around $TWR_B$ is dominated by *Taxodium distichum* seedlings that were approximately 1 m tall in 2008, with a dense strip of *Typha latifolia* plants located about 25 m east of the tower. The water table depth around $TWR_B$ was greater than 30 cm in 2008, and was slightly higher in 2009. A pumping system used to drain the site when it was operated as a farmland still exists and is functional, and the pumps were turned on from August 18, 2008 – Sept. 1, 2008 to drain the area around $TWR_B$ as part of an artificial drawdown experiment.

The eddy covariance and meteorological instrumentation mounted on the $MMS$ was also situated in two vegetatively and hydrologically distinct areas of the wetland over the course during the study. In 2008, the $MMS$ was parked on a dryland portion of the wetland similar to the area surrounding $TWR_A$ (Figure 6.1) and with water table depth well below the surface. In 2009, the $MMS$ was parked on an elevated road surrounded on each side by flooded areas of the $TRP$ populated with a range of native shrubs, trees, and grasses. The average water table depth around the $MMS$ in 2009 was over 30 cm.

Eddy covariance and meteorological measurements were collected from 8/11/2008 to 9/29/2008, and from 6/4/2009 to 9/2/2009. Gas profile measurements at $TWR_A$ and $TWR_B$ were made on 12 days in 2008 (8/16 to 8/21 and 8/31 to 9/5). Of those 12 measurement days, 6 were within the period of the artificial drawdown experiment. Gas profile measurements were also made on 12 days in 2009 (7/8 to 7/10, 7/21 to 7/23, 8/11 to 8/13, and 8/31 to 9/2).

6.3.2 Eddy covariance measurements - Set-up and quality control

Sensible heat flux ($H$), latent heat flux ($LE$), and net ecosystem CO$_2$ flux ($F_{CO2,EC}$) were measured at the $MMS$ continuously during the study period using an eddy covariance system comprised of triaxial sonic anemometer (CSAT3, Campbell Sci-
Figure 6.1: The Timberlake Restoration Project. (a) shows LiDAR elevation data and the locations of the atmospheric measurement stations. The concentration profile towers are marked with circles, and the Mobile Micro-meteorological Station (MMS) locations are marked with squares for 2008 (northern location) and 2009 (southern location). (b) shows the area around $TWR_B$ during an artificial drawdown experiment. (c) shows the area around $TWR_A$, and (d) shows the MMS during the 2008 deployment.
entific, Logan, UT, USA) and open-path infrared gas analyzers (IRGA, LI-7500, Li-Cor, Lincoln, NE, USA). The eddy covariance instrumentats were mounted on a telescoping mast atop the roof of a full size cargo van and extended to a height of 9 m above ground. The mast was leveled and secured using a guy wires and jacks on the underside of the MMS. Measurements of vertical wind velocity and scalar concentrations were collected and stored at 10 Hz using CR-3000 dataloggers (Campbell Scientific). Flow statistics and fluxes were processed offline using MATLAB into hourly averages. A 3-D coordinate axis rotation was applied within each averaging period to ensure that vertical wind measurements were oriented perpendicular to the mean wind flow as described in Wilczak et al. (2001). The Webb–Pearman–Leuning correction (Webb et al., 1980) for the effects of air density fluctuations on CO$_2$ and H$_2$O flux measurements was applied. Measured latent heat fluxes were converted to ET, in units of mm per hour, using the temperature-dependent latent heat of vaporization for water vapor.

EC flux measurements were subjected to a quality control procedure to remove excessive sensor noise and data collected under very stable conditions. First, spikes in the data record were removed using the algorithm recommended for FLUXNET (Papale et al., 2006). Fluxes were then filtered to remove data collected during stable atmospheric conditions using the friction velocity ($u^*$) filtering method described in Chapter 2, Section 3.5. Finally, fluxes were rejected if the half-hourly flux magnitude was outside the logical expectation window of -200 to 1000 W m$^{-2}$ for $LE$ and $H$, and -1.0 to 1.0 mg CO$_2$C m$^{-2}$ s$^{-1}$ for $F_{CO2,EC}$.

6.3.3 Meteorological and edaphic measurements

A suite of meteorological measurements was also collected at the study site. At the MMS locations, $R_n$ was measured with a CNR1 net radiometer (Kipp & Zonen, Delft, the Netherlands). Photosynthetically active radiation (PAR) was measured
with a LI-190SA quantum sensor (Li-Cor). \( T_a \) and \( RH \) were measured with HMP45C Ta/RH (Vaisala, Finland) probes positioned at a height of 1.8 m near the MMS, and then used to calculate vapor pressure deficit \( (D) \). Wind direction \( (WD) \), \( \bar{u} \), and \( u^* \) were measured from the velocity components of the sonic anemometer. A vertical profile of mean air temperature was also measured on the MMS (at heights of 1.8 m, 3.5 m, 4.5 m, and 5.5 m) using thermocouples. All meteorological variables were sampled at frequencies ranging from 0.1 to 10 Hz and then averaged to hourly values.

Soil moisture content was measured at three depths at both \( TWR_A \) and \( TWR_B \) (5, 10, and 15 cm) using Type ML1 ThetaProbe soil moisture sensors (Delta-T Devices, Cambridge, UK). Soil temperature was measured at the same locations and depths with thermocouples. Ground heat flux \( (G_h) \) was measured at two locations at a depth of 5 cm at both \( TWR_A \) and \( TWR_B \) using HFT-3 soil heat flux plates (REBS, Campbell Scientific, Inc.). When the measurement of any meteorological variable was unavailable due to equipment failure or other error, a continuous record was obtained by fitting a linear regression between measurements from the sensor of interest and data from a similar instrument located within the study site or, if no such instrument was available, from a similar instrument situated at a weather station in Plymouth, N.C. (about 30 miles west of the study site). The regression equation was then used to fill the missing data.

6.3.4 Vertical gas concentration measurements and quality control

At \( TWR_A \) and \( TWR_B \), the vertical concentration gradients of CO\(_2\), CH\(_4\), and N\(_2\)O were measured at hourly intervals from 0800 to 1700 hours on sampling days in 2008, and from 0900 to 1700 on sampling days in 2009. Occasionally, weather events or equipment failures prevented the collection of the mean concentration gradients. Air was pumped from inlet locations at heights of 2 m, 3.2 m, 4 m, and 5 m from a PVC tower using Watson Marlow peristaltic pumps (Model 504S) into airtight
PFTE sampling bags. The sampling bags were changed every hour and two 10-ml subsamples (subsample A and subsample B) were created from each sampling bag for analysis in the laboratory with a gas chromatographer (GC). The GC was calibrated using standard gases with known concentrations in ppmv. The standard gases were changed for the 2009 measurement campaign in an effort to reduce measurement error by selecting standards that better matched the observed atmospheric concentration range of each gas. Measured concentrations were converted to ppm from ppmv using air pressure measurements from the Li-7500 and temperature measurements made with thermocouples located next to each inlet port on the sampling towers.

Gas concentrations were determined for subsamples A and B within 10 days of the field sampling date. No two subsamples were analyzed in the same GC batch to prevent data loss from GC malfunction. The relative error between subsample A and subsample B was determined for each tower for each day for each gas. When that error exceeded 10%, the concentration from the subsample that was closest to measured concentration from the same inlet location at the same time on either a day directly preceding or following the measurement date was selected as the best concentration estimate. Otherwise, concentrations from subsample A and B were averaged.

The concentration profile data were used to determine the $SG$ flux estimates as described in Section 6.2.2. Flux estimates were removed if they exceeded the thresholds for each gas shown in Table 6.1. These thresholds were generally set to remove the extreme 5% of data from the flux records. Additionally, flux estimates were rejected if the absolute value of $H$ was less than 10 W m$^{-2}$, as the matching of $H$ to $\partial T_a/\partial z$ is difficult for low magnitude fluxes and small temperature gradients.
<table>
<thead>
<tr>
<th>Flux</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{CO2,SG}$ (mg CO$_2$-C m$^{-2}$ h$^{-1}$)</td>
<td>-5000</td>
<td>4000</td>
</tr>
<tr>
<td>$F_{CH4,SG}$ (mg CH$_4$-C m$^{-2}$ h$^{-1}$)</td>
<td>-100</td>
<td>100</td>
</tr>
<tr>
<td>$F_{N2O,SG}$ (mg N$_2$O-N m$^{-2}$ h$^{-1}$)</td>
<td>-60</td>
<td>60</td>
</tr>
</tbody>
</table>

Table 6.1: The thresholds used with the scalar gradient methodology to filter spurious estimates.

6.3.5 Error estimation

Eddy covariance estimates of surface fluxes represent observations of the true flux modified by two types of error: random measurement error ($\delta$) and systematic error (Richardson et al., 2006). The former term incorporates all instrument errors, as well as errors due to surface heterogeneity and stochastic noise inherent to turbulent transport. The latter term incorporates persistent biases associated with experimental design and methodology, which may be best assessed via independent flux measurements. Here, the former error term is calculated by determining the mean absolute error ($\sigma(\delta)$) in $LE$ and $F_{CO2,EC}$ by comparing acceptable flux measurements collected during the same time period on consecutive days under similar environmental conditions ($u$ within 1 m/s, $D$ within 0.5 kPa, and $R_n$ within 50 W m$^{-2}$) as discussed in Chapter 2, Section 6.1.

Error in the estimates of the GHG fluxes estimated using the $SG$ approach has three principle sources:

1. Measurement error in $H$. The sensible heat flux is used directly to estimate the $SG$ fluxes via Equation 6.3. Errors in $H$ can be large (i.e. 10 to 20%), but relative errors should be smaller during the daytime periods when the concentration profile measurements were made (Wilson and Baldocchi, 2000). To an extent, the magnitude of this error may be assessed by assessing energy balance closure.
2. *Methodological errors.* Methodological errors are introduced when the assumptions of uniform $K_h$ for all scalars is violated, or when a mismatch exists between $H$ and the temperature profile $\partial T_a/\partial z$. To an extent, the latter may be assessed by comparing inferred diffusivities (i.e. $K_{h,\text{meas}} = (H/\rho c_p) (\partial T_a/\partial z)$) to modeled diffusivities following similarity theory. These comparisons are made in the discussion section.

3. *Concentration measurement errors.* The SG method is sensitive to errors in concentration measurements from the GC instrumentation, particularly since variation in GHG concentrations is small over the narrow vertical range of the concentration profiles (i.e. $<5$ m). The magnitude of this error is assessed directly by repeating flux calculations using only subsample A concentration estimates, and again using only subsample B concentration estimates, for periods for which reliable measurements from both subsamples are available. The SG method is further subject to errors related to temperature profile measurements; however, these errors will likely be small relative to the concentration profile measurement errors.

### 6.4 Results

#### 6.4.1 Meteorological conditions

Daytime air temperatures and relative humidity were high during the study period (Figure 6.2), as is typical for eastern N.C. in late summer. The average $T_a$ and $RH$ for daytime periods when gas concentration profile measurements were taken was remarkably similar in 2008 and 2009 ($\langle T_a \rangle = 26.6$ and 26.8 °C, and $\langle RH \rangle = 68$ and 69 % for 2008 and 2009, respectively), though day-to-day variation in these variables is considerable (Figure 6.2). Average $R_n$ was much lower for these same periods in 2009 as compared to 2008 ($\langle R_n \rangle = 320$ vs 422 W m$^{-2}$, respectively) due primarily
Figure 6.2: Meteorological conditions at the site for the days on which concentration profile measurements were conducted. Air temperature ($T_a$) is shown in (a) and (b) for 2008 and 2009, respectively. Relative humidity ($RH$) is shown in (c) and (d), and net radiation ($R_n$) is shown in (e) and (f). The thick gray bars represent the average value for the entire study period in each year. The vertical dotted lines denote breaks in the x-axis time series.

to a series of five cloudy days during the latter half of the 2009 study period (Figure 6.2f).

Precipitation events on gas concentration collection days were generally limited to afternoon convective rainfall events that were brief in duration. However, nearly 7 cm of rain fell at the TRP during the drawdown experiment between August 26-30 (DOY 239-242), 2008, which were the four days directly preceding the second set of gas concentration measurements. This large rain event combined with the high water
retention capacity of the poorly drained, mucky soil caused the soil around $TWR_B$ to remain extremely wet for the duration of the drawdown experiment. Around $TWR_B$, soil moisture at 5 cm depth declined to about 70-80% of the saturation value before the large rain event; however, the impact of the drawdown experiment was not detectable on soil moisture measured at 15 cm. The drawdown had no measureable impact on soil moisture at any depth in the area surrounding $TWR_A$.

6.4.2 Data availability

Hourly-averaged eddy covariance data (i.e. $F_{CO2,EC}$, $ET$, $H$) were collected for 985 hours during the 2008 study period, and for 762 hours during the 2009 study period. After the quality control filters described in Section 6.3.2 were employed, 67% of the 2008 and 56% of the 2009 MMS flux data remained, though many of the filtered data points were collected at night. For periods when $PAR > 100 \ \mu mol \ m^{-2} \ s^{-1}$, 80% and 73% of the data remained.

$SG$ flux estimates at $TWR_A$ and $TWR_B$ were available for a total of 116 hours in 2008 and 96 hours in 2009. Of these hourly flux estimates, 1% were rejected for low $H$ in 2008, and 16% were rejected for low $H$ in the more cloudy 2009 growing season. An additional 6-7% of the data were rejected for each gas due to pump or gas collection bag failures. And, by design, an average of 5% of the data were filtered for falling outside of the thresholds given in Table 6.1. After these two filters were applied, 79% of the data remained across the range of greenhouse gas fluxes that were measured (Table 6.2). In 2008, the atmospheric concentration of $N_2O$ around both $TWR_A$ and $TWR_B$ fell below the lowest available $GC$ standard calibration gas; therefore, $N_2O$ fluxes could not be determined in 2008. In 2009, lower standard gases were acquired and $N_2O$ fluxes were estimated.
Flux 2008 2009 2008 2009

\[ F_{\text{CO}_2,SG} \text{ (mg CO}_2\text{-C m}^{-2}\text{h}^{-1}) \]
84 71 84 74

\[ F_{\text{CH}_4,SG} \text{ (mg CH}_4\text{-C m}^{-2}\text{h}^{-1}) \]
86 75 85 75

\[ F_{\text{N}_2\text{O,SG}} \text{ (mg N}_2\text{O-N m}^{-2}\text{h}^{-1}) \]
75 76

Table 6.2: Percentage of measured flux data remaining after the application of quality control procedures for each measurement location. Fluxes of \( \text{N}_2\text{O} \) were not estimated in 2008. Data availability is less in 2009 due to lower sensible heat flux.

6.4.3 Footprint considerations

A scalar flux footprint model (Hsieh et al., 2000) was applied with measured wind and atmospheric stability conditions from 2009 for \( z = 10 \text{ m} \), which is slightly higher than the measurement height of the eddy covariance instrumentation on the \( \text{MMS} \).

The average location of the peak of the footprint source weight function with respect to tower location \( (x_p) \) is shown in Figure 6.3.

For a 10 m measurement height, \( x_p \) is less than 100 m 88% of the time, and is less than 50 m over 80% of the time (Figure 6.3b). These dimensions are well within the fetch of the representative vegetation areas around the \( \text{MMS} \) in both 2008 and 2009 (see Figure 6.1). Footprints for concentration measurements (like those measured at \( \text{TWR}_A \) and \( \text{TWR}_B \)) are generally larger than flux footprints; however, concentration footprints decline considerably at lower measurement heights like those associated with the concentration profiles (i.e. 2, 3.2, 4, and 5 m). Hence, the area of the concentration profile footprint is likely to be comparable or smaller than the \( \text{MMS} \) flux footprint and to largely lie within the characteristic vegetation surrounding \( \text{TWR}_A \) and \( \text{TWR}_B \).

Additionally, I note that the wind tides experienced at the site cause wind direction to be primarily from the NNW-NNE and SSE-SSW (Figure 6.3a). For these
Figure 6.3: Flux footprint considerations for the MMS. (a) shows a histogram of wind direction (\(WD\)), with pronounced wind tides that tend to limit wind direction to the NNW to NNE or SSW to SSE. The location of the peak of the source weight function (\(x_p\)) during the 2009 study period with respect to the measurement location (black square) as estimated using the model of Hsieh et al. (2000) is given in (b) for a 10 m measurement height.

Because a detailed, high-resolution spatial characterization of the vegetation at the site does not exist, the footprint model was not used to filter data as part of the quality control procedures. However, the results from this simulation exercise suggest that most of the time, the flux footprints from atmospheric measurement locations lie well within the ecosystem dimensions.
6.4.4 CO$_2$ and energy fluxes measured on the MMS

The area around the MMS was a sink of CO$_2$ during daytime periods and a source of CO$_2$ at night in both 2008 and 2009 (Figure 6.4). The data were fitted to the widely used non-rectangular hyperbola model (see Equation 4.2). The parameters $\alpha$, $\beta$, and $\gamma$ were all larger for the 2008 data ($0.72$ mg C $\mu$mol$^{-1}$, $884$ mg C m$^{-2}$ h$^{-1}$, and $104$ mg C m$^{-2}$ h$^{-1}$), respectively, as compared to 2009 ($0.36$ mg C $\mu$mol$^{-1}$, $362$ mg C m$^{-2}$ h$^{-1}$, and $87$ mg C m$^{-2}$ h$^{-1}$). Hence, both assimilation capacity and ecosystem respiration were higher in magnitude in the dryland area that surrounded the MMS in 2008 as compared to the inundated area in which the van was situated in 2009. The curvature parameter $\theta_f$ was lower in 2008 than 2009 (0.2 vs 0.70, respectively) as the flux data saturated more quickly with light in the upland area (Figure 6.4).

Latent heat flux was similar for daytime periods (0900 - 1800) with high light ($PAR > 1000$ $\mu$mol m$^{-2}$ s$^{-1}$) in 2008 and 2009, with average values of 162 W m$^{-2}$ in 2008 and 176 W m$^{-2}$ in 2009. However, the diurnal patterns of $LE$ differ between years, with more late-afternoon $LE$ in 2009 vs 2008 (Figure 6.5).

The ratio of $LE:R_n$ and $H:R_n$ are also shown in Figure 6.5. Both of these ratios decreased in 2009 versus 2008, suggesting that a significant portion of the available energy was directed towards heating the standing water surrounding the 2009 MMS location. This ground heat flux ($G_h$) was not measured around the MMS in 2009. $G_h$ was measured around TWR$_A$ in 2008 in vegetation that is very similar to the vegetation surrounding the MMS in 2008, and the sum of $LE + H + G_h$ for 2008 is also shown in Figure 6.5 as a fraction of $R_n$. During daytime periods, this energy balance ratio closure averaged 85%, which is comparable to energy balance closure rates at other eddy covariance sites (Wilson et al., 2000).
Figure 6.4: The relationship between eddy covariance derived CO\textsubscript{2} flux (\(F_{\text{CO}_2,\text{EC}}\)) and photosynthetically active radiation (\(PAR\)) using 2008 and 2009 Mobile Micro-meteorological Station (MMS) data. Negative fluxes indicate assimilation. Symbols denote averages binned by light and error bars indicated standard errors of the mean. The smoothed curves are the non-rectangular hyperbolic model of Eq. 6.6 parameterized by the data using non-linear regression.

6.4.5 GHG flux estimates using the SG method

The SG-derived GHG flux estimates, averaged for the 2008 and 2009 study periods, are shown in Figure 6.6. The area around TWR\textsubscript{A} was a sink for all three GHGs throughout the study period, though the magnitude of the sink for CO\textsubscript{2} and CH\textsubscript{4} decreased in 2009 when the area was under several inches of standing water (\(\langle F_{\text{CO}_2,\text{SG}}\rangle = -366 \text{ and } -195 \text{ mg } \text{CO}_2\text{-C m}^{-2} \text{ h}^{-1}\) in 2008 and 2009, respectively, and \(\langle F_{\text{CH}_4,\text{SG}}\rangle = -7.7 \text{ and } -4.2 \text{ mg } \text{CH}_4\text{-C m}^{-2} \text{ h}^{-1}\) in 2008 and 2009, respectively). The flux estimates
Figure 6.5: Eddy-covariance derived energy fluxes from the Mobile Micro-meteorological Station MMS averaged over time of day. Latent heat (LE) flux during high light conditions is shown for both study years in (a). The ratio of LE to net radiation ($R_n$) and the ratio of sensible heat flux ($H$) to $R_n$ are shown in (b) and (c), respectively. Mean radiation balance closure, determined using 2008 LE, H, and $R_n$ measurements from the Mobile Meteorological Station and 2008 ground heat flux measurements ($G_h$) from the area around TWR_A, is shown in (d).

suggest the area around TWR_B, which is typically under >30 cm of standing water, was a marginal though insignificant sink of CO$_2$ in 2008 ($\langle F_{CO2,SG} \rangle = -89 \pm 130$ mg CO$_2$-C m$^{-2}$ h$^{-1}$) and a slightly stronger sink in 2009 ($\langle F_{CO2,SG} \rangle = -178 \pm 169$ mg CO$_2$-C m$^{-2}$ h$^{-1}$). This area was an insignificant source of CH$_4$ in 2008 ($\langle F_{CH4,SG} \rangle = 1.3 \pm 3.5$ mg CH$_4$-C m$^{-2}$ h$^{-1}$, Figure 6.6) and a stronger and significant source in 2009 ($\langle F_{CH4,SG} \rangle = 5.2 \pm 2.2$ mg CH$_4$-C m$^{-2}$ h$^{-1}$). The error associated with the
flux estimates is large and in many cases the confidence intervals (determined from the standard error of the mean) overlap across locations and between years.

As mentioned in Section 6.4.2, N\textsubscript{2}O fluxes were not calculated in 2008 due to a mismatch between atmospheric gas concentrations and the standard gases for the GC. In 2009, the SG method suggests the area around TWR\textsubscript{A} was an insignificant sink for N\textsubscript{2}O in 2008 (\(\langle F\textsubscript{N2O,SG} \rangle = -0.56 \pm 0.67 \text{ mg N}_2\text{O-N m}^{-2} \text{ h}^{-1}\)) and the area around TWR\textsubscript{B} was a significant source for N\textsubscript{2}O in 2009 (\(\langle F\textsubscript{N2O,SG} \rangle = 1.21 \pm 0.83 \text{ mg N}_2\text{O-N m}^{-2} \text{ h}^{-1}\)).

Across all methods, locations, and years, the magnitude of CO\textsubscript{2} fluxes is roughly an order of magnitude higher than CH\textsubscript{4} fluxes, which in turn is roughly an order of magnitude higher than N\textsubscript{2}O fluxes. Fluxes from inundated locations were generally more positive than fluxes from dry locations, and fluxes from the permanently inundated area around TWR\textsubscript{B} were generally more positive than fluxes from the occasionally inundated area around TWR\textsubscript{A}, even when the latter area was under water.

6.4.6 Drawdown impact

Daily averaged CO\textsubscript{2} and CH\textsubscript{4} fluxes for the area around TWR\textsubscript{B} in 2008 are shown in Figure 6.7. The two lowest daily-averaged flux estimates for CO\textsubscript{2} were measured during the drawdown period, and on average, \(F\textsubscript{CO2,SG}\) is significantly lower during the drained period (\(\langle F\textsubscript{CO2,SG} \rangle = -314 \pm 210 \text{ mg CO}_2\text{-C m}^{-2} \text{ h}^{-1}\)) versus the inundated period in 2008 (\(\langle F\textsubscript{CO2,SG} \rangle = 132 \pm 150 \text{ mg CO}_2\text{-C m}^{-2} \text{ h}^{-1}\)). Methane fluxes increased somewhat in dry versus inundated periods (\(\langle F\textsubscript{CH4,SG} \rangle = 2.12 \pm 6.4\) and \(0.54 \pm 3.0 \text{ mg CH}_4\text{-C m}^{-2} \text{ h}^{-1}\), respectively) though the difference is not significant. The effect of the drawdown was not explicitly assessed at TWR\textsubscript{A} because the manipulation experiment produced no measurable impact on soil moisture at TWR\textsubscript{A}.  

165
6.4.7 Error in the flux estimates

Random measurement error for the $EC$ fluxes was determined using the paired-observation approach described in Chapter 2. For $F_{CO2,EC}$, $\sigma(\delta) = 0.049$ and 0.041 mg CO$_2$-C m$^{-2}$ h$^{-1}$ in 2008 and 2009, respectively, which is within the range of growing season estimates for the same quantity reported in the Richardson study for seven eddy covariance sites (0.0072 to 0.52 mg CO$_2$-C m$^{-2}$ h$^{-1}$). The random measurement error for $LE$ was 76 W m$^{-2}$ for data collected when $R_n > 400$ W m$^{-2}$, which is also
Figure 6.7: The impact of the artificial drawdown experiment on CO$_2$ and CH$_4$ fluxes from the area around TWR$_B$ in 2008. Symbols show daily averaged flux rates and error bars show the standard error of the mean. The shaded area indicated days on which the standing water around TWR$_B$ was removed by pumping. The vertical dashed lines indicates an x-axis discontinuity.

well within the range of errors for similar data from other sites.

Error in eddy covariance datasets may also be evaluated by assessing energy balance closure, which was acceptable at this site (Figure 6.5). On average, the sum of $LE$, $H$, and $G$ is within 15% of measured $R_n$ for the daytime portions of the study period for which acceptable measurements of all three fluxes are available. This level of closure is within the average radiation balance closure of 20% reported in a synthesis of data from a range of FLUXNET sites (Wilson and Baldocchi, 2000).

The error in the $SG$ estimates of $F_{CO2}$, $F_{CH4}$, and $F_{N2O}$ could not be assessed
using the paired-point approach due to the low number of available data. Instead, it was assessed via more indirect analyses. First, I note that the error bars shown in Figures 6.6 and 6.7 (which represent the standard error of the mean) are quite large, typically on the order of the flux magnitudes themselves, indicating large uncertainty associated with this method.

The extent to which this uncertainty could be attributed to error associated with the GC concentration measurements was assessed by comparing flux estimates determined using gas concentration data from subsample A with those flux estimates determined using gas concentrations from subsample B. Theoretically, these two flux estimates should be the same, and any differences reflect errors associated with the GC instrument. At the hourly time scale, these estimates are not well correlated, though when averaged over each seasonal study period they are qualitatively similar (Figure 6.8). However, differences exist, and in some cases the error between the two estimates is comparable to the magnitude of the flux (for example, see the \( F_{CH4} \) estimates from \( TWR_A \) in 2009).

6.5 Discussion

The range of GHG fluxes from wetland ecosystems reported in the literature is large (see Table 1.1). Both source and sink fluxes of \( CO_2 \), \( CH_4 \), and \( N_2O \) have been measured, and in the case of \( CH_4 \) and \( N_2O \), flux estimates spans several orders of magnitude across ecosystems. In this study, I explored variability with respect to water table depth on the ecosystem exchange of these gases. Results from two micro-meteorological techniques suggest that the fluxes of \( CO_2 \), \( CH_4 \), and \( N_2O \) in a large, restored temperate wetland all tend to become more positive (i.e. less of a sink) in the presence of standing water (Figure 6.6). The results also suggest a significant daytime sink of \( CO_2 \) and \( CH_4 \) in dryland portions of the wetland during the study period, and a source of \( CH_4 \) and \( N_2O \) in a permanently inundated portion of the wetland.
Figure 6.8: Scalar gradient flux estimates derived from subsample A (filled symbols) and subsample B (open symbols). In the absence of any concentration measurement error, the two paired flux estimates should be equivalent. Squares indicate dryland fluxes and circles indicate fluxes from inundated areas. Symbols show mean values and error bars show the standard error of the mean.

In the following sections, the variability and uncertainty of these flux estimates is discussed in more detail, and I conclude by making some recommendations for future applications of these methodologies.

6.5.1 CO$_2$ fluxes

Both the EC and SG methodologies suggest a daytime sink for CO$_2$ for all of the study locations, though the magnitude of this sink varies. Following thermodynamic predictions, the sink strength is largest in the absence of standing water (i.e. the area
around $TWR_A$ and the $MMS$ in 2008, $\langle F_{CO2} \rangle = -367$ to -407 mg CO$_2$-C m$^{-2}$ h$^{-1}$), and is smallest in areas with water table depths well above the surface (i.e. the area around $TWR_B$ in 2008 and 2009, and the area around the $MMS$ in 2009, $\langle F_{CO2} \rangle = -170$ to -89 CO$_2$-C m$^{-2}$ h$^{-1}$). In addition to hydrologic state, vegetation cover should have a large impact on net ecosystem CO$_2$ fluxes by controlling photosynthetic carbon assimilation. Vegetation around $TWR_A$ and the 2008 $MMS$ location was similar and comprised of dense grasses and shrubs, which may explain the why flux estimates from these areas are so well conserved in 2008 across methods (i.e. Figure 6.6). Vegetation was more sparse in the areas around $TWR_B$ and the 2009 $MMS$ location and was more heavily dominated by tree seedlings, which differed in species composition between these two flooded locations. All of these flux estimates are within the range of daytime net ecosystem CO$_2$ flux previously measured over other wetland ecosystems (Table 1.1).

The area around $TWR_A$ was dry in 2008 and flooded in 2009, permitting an investigation of the impact of inundation alone on $F_{CO2}$, as vegetation cover at this location was similar in both years. While this area was estimated to be a sink in both years, fluxes were 50% lower in 2009 as compared to 2008 (Figure 6.6). Inundation should suppress heterotrophic respiration, which alone should produce more negative CO$_2$ fluxes. On the other hand, standing water also inhibits photosynthesis, which would promote more positive fluxes. Therefore, the less negative fluxes in 2009 could represent greater inhibition of photosynthesis versus respiration under standing water. However, solar radiation was considerably less in 2009 (Figure 6.2), and limited photosynthesis in that year (see Figure 6.4).

The drawdown experiment around $TWR_B$ also permits an exploration of the impact of water table variation on GHG fluxes over constant vegetation. I predicted that the drawdown would promote more positive CO$_2$ fluxes as the removal of standing water should promote heterotrophic respiration. Instead, significantly more
negative CO$_2$ fluxes were observed during the drawdown phase ($\langle F_{CO2,SG} = -314 \pm 210$ mg CO$_2$-C m$^{-2}$ h$^{-1}$) as compared to the fluxes observed when the area was under standing water in 2008 ($\langle F_{CO2,SG} = 132 \pm 150$ mg CO$_2$-C m$^{-2}$ h$^{-1}$). Since the soil around $TWR_B$ remained nearly saturated for the duration of drawdown experiment, the absence of a clear increase in heterotrophic respiration during the drawdown phase is quite possible. However, the strong increase in sink strength during the drawdown period is surprising given the relatively sparse vegetation cover around $TWR_B$. The removal of standing water could promote photosynthesis (which is typically decreased in flood areas), but it seems unlikely that the drawdown would enhance photosynthesis but not soil respiration.

When the drawdown period is excluded from the analysis, the area around $TWR_B$ switched from a source of CO$_2$ in 2008 to a sink in 2009. As the area around $TWR_B$ is dominated by young and rapidly growing tree seedlings, part of this shift may be explained by greater leaf area in 2009.

Finally, I note that all measured areas of the $TRP$ were net daytime CO$_2$ sinks during the study period. However, the same can not be assumed for all time periods or on an annual basis, given that the site is a nocturnal source of CO$_2$. (Figure 6.4), and that sink strength is likely to be highest in the peak of the growing season (i.e. the study period) when leaf area and day length at a maximum.

6.5.2 CH$_4$ fluxes

The estimates of $F_{CH4}$ were roughly an order of magnitude smaller than $F_{CO2}$. Following theoretical predictions, $F_{CH4,SG}$ was negative for the dry area around $TWR_A$ in 2008 ($\langle F_{CH4,SG} = -7.7 \pm 3.5$ mg CH$_4$-C m$^{-2}$ h$^{-1}$) indicating the presence of methanotrophic bacteria. These flux magnitudes are among the lowest reported in the literature synthesis of Table 1.1. The sink strength for methane around $TWR_A$ decreased in the presence of standing water in 2009 (to $\langle F_{CH4,SG} = -4.2 \pm 2.5$ mg
CH$_4$-C m$^{-2}$ h$^{-1}$).

The area around TWR$_B$ was nearly neutral with respect to CH$_4$ in 2008 and was a significant source of CH$_4$ in 2009 (Figure 6.6), with flux magnitudes well within the range of values previously reported in the literature (Table 1.1). In 2008, the drawdown had no detectable impact on CH$_4$ fluxes, which is consistent with negligible changes in soil moisture content during the drawdown. The drawdown impact, therefore, can not explain the large difference between 2008 and 2009 methane fluxes around TWR$_B$. Rather, I propose that this difference is related to varying sulfate concentration, which has been previously shown to inhibit methanogenesis (Dodla et al., 2009). In 2008, this portion of the TRP experienced significant salt water intrusion with average SO$_4^-$ concentrations in the area near TWR$_B$ of 210 mg/L during the 2008 study period. In 2009, the brackish water receded and mean study period SO$_4^-$ concentrations declined to 10 mg/L, coincident with the higher CH$_4$ effluxes.

6.5.3 N$_2$O fluxes

$F_{N2O}$ was not estimated in 2008 when atmospheric concentrations were below the detections limits for the GC. New calibration gases permitted N$_2$O fluxes to be to be estimated in 2009. The SG method estimates a N$_2$O sink around TWR$_A$ in 2009 ($\langle F_{N2O}\rangle = -0.55 \pm 0.67$ mg N$_2$O m$^{-2}$ h$^{-1}$) and a N$_2$O source from the deeper standing water around TWR$_B$ ($\langle F_{N2O}\rangle = 1.21 \pm 0.83$ mg N$_2$O m$^{-2}$ h$^{-1}$). These flux magnitudes are roughly an order of magnitude smaller than the methane fluxes and two orders of magnitude smaller than the CO$_2$ fluxes. The source fluxes from around TWR$_B$ are expected from anoxic soils under standing water and are well within the range of reported values from other ecosystems (Table 1.1).

The N$_2$O sink measured around TWR$_A$ in 2008 is significantly lower than most of the N$_2$O fluxes recently reported in the literature (Table 1.1), though negative fluxes of similar magnitude have been previously reported in older studies (Longoria-
Ramirez et al., 2003; Merbach et al., 1996). The magnitude of this flux does not appear to be a result of concentration measurement error (see Figure 6.8). N₂O is consumed by soil microbes during the final phase of denitrification, when nitrous oxide is reduced to N₂. Typically, nitrous oxide during the early phases of denitrification produced at a much higher rate during the earlier stages of denitrification, leading to a net positive efflux. When N₂O diffusion through the soil matrix is limited, or when nitrate availability is low, net N₂O consumption can occur (Chapuis-Lardy et al., 2007). Hence, a mechanism exists to explain the significantly negative fluxes observed in this study, though the magnitude of the flux around TWR₄ is nonetheless surprising and warrants further process-based investigation.

6.5.4 H₂O fluxes

Water vapor differs from the other greenhouse gases measured here in that water vapor does not contribute to long-term climate change given its short atmospheric residence time. However, evapotranspiration is nonetheless important to quantify given the close coupling of transpiration and photosynthesis, as well as the link between ET and the humidity concentration in the boundary layer known to impact rainfall triggering (Juang et al., 2007; Siqueira et al., 2009). Here, ET could not be measured with the SG method, but was estimated with the eddy covariance instrumentation over the 2008 and 2009 MMS locations. Surprisingly, the magnitude of mean daytime ET was similar over both locations, despite significant differences in water table level and vegetation (⟨ ET ⟩ = 0.24 and 0.25 mm h⁻¹ in the 2008 versus 2009 locations). The pattern of LE fluxes with time of day differs noticeably (Figure 6.5), with more mid-to-late afternoon LE observed over the dry location. Because plant stomata close in afternoon periods to prevent excessive water loss under increasing vapor pressure deficit, the afternoon reduction in the 2008 LE record may suggest a greater contribution of transpiration to ET in 2008 versus 2009.
Figure 6.9: The relationship between surface conductance ($G_T$) and vapor pressure deficit ($D$) for the dryland area around (a) the mobile meteorological station (MMS) in 2008 and (b) the inundated area around the MMS in 2009. Also shown is the equation $G_T = m \ln(D) + b$ parameterized with parameters determined from linear regression with data from each year. If evapotranspiration is dominated by transpiration, then the slope of the line should be in the range $m = 0.55 - 0.65$.

To further assess the relative contribution of transpiration versus evaporation to total $ET$, total surface conductance ($G_T$) was estimated from the $LE$ records. According to physiological theory, the response of $G_T$ to $D$ should follow a log-linear relationship with a slope of $m \sim 0.55$ to $0.65$ mmol m$^{-2}$ s$^{-1}$ kPa$^{-1}$ (Katul et al., 2009; Oren et al., 1999b) if $ET$ is dominated by transpiration. In 2008, when the MMS was positioned in a dryland portion of the wetland, the slope of this relationship was $m = 0.6$ (Figure 6.9a). However, in 2009, when the MMS was positioned in an inundated
portion of the wetland, the slope was significantly decreased \((m = 0.26, \text{Figure 6.9b})\), suggesting a larger contribution from evaporation which compensates for decreased transpiration in the flooded location. Since transpiration and assimilation are tightly coupled, this result is also supported by the decreased assimilation rates observed around the 2009 MMS location as compared to 2008 (Figure 6.4).

### 6.5.5 Flux estimation error

The sources of error for eddy covariance measurements are well known and have been discussed in this text and in the literature (Baldocchi, 2003; Friend et al., 2007; Novick et al., 2009b; Oren et al., 2006; Stoy et al., 2006b). They can be on the order of 10-20% for the full annual cycle, though they tend to be relatively smaller during the daytime conditions that comprise the majority of the study period. Because the estimates of measurement error (i.e. \(\sigma(\delta)\)) for both \(F_{CO2,EC}\) and \(LE\) are comparable to those reported from other sites (see Section 6.4.7), and because the energy balance closure is above 85% for most daytime periods, there is no reason to suspect that the eddy covariance records from the MMS are infected with errors beyond those which are typically associated with this methodology.

As the error bars in Figures 6.6 to 6.8 suggest, the uncertainty associated with the scalar gradient is quite large, with estimates for the standard error of the mean that are often at times on the same order as the flux magnitudes themselves. Much of this error may be attributable to GC concentration measurement error, which is as high as 3.5% in this study (Table 6.3) and which propagates to large errors in flux estimation as the vertical concentration gradients are small. An exercise that estimated flux rates using subsample A and subsample B produced estimates for all gases that were qualitatively similar but with magnitudes that differed considerably, particularly for smaller magnitude fluxes (see Figure 6.8).

To further explore the uncertainty attributable to concentration measurement
Table 6.3: The gas chromatographer concentration measurement error for CO$_2$, CH$_4$, and N$_2$O, determined as the mean relative error between subsample A and subsample B, which should be equal in the absence of any measurement error. Fluxes of N$_2$O were not estimated in 2008.

Table 6.3: The gas chromatographer concentration measurement error for CO$_2$, CH$_4$, and N$_2$O, determined as the mean relative error between subsample A and subsample B, which should be equal in the absence of any measurement error. Fluxes of N$_2$O were not estimated in 2008.

error, I conducted a simulation exercise designed to determine how concentration measurement error propagates to flux estimation error (see Appendix A) as a function of data record length. This simulation exercise suggests that, on average, at least 100 data points are needed to recover the mean ‘true flux’ to within 10% if the concentration error is 1%. Relative errors near 20% persist even for data record that are over 100 points long for 2% and 4% concentration errors. The large data requirement illustrated by this exercise may explain, in part, the absence of any significant trends between SG flux estimates and meteorological variables, as the data records themselves are only $\sim$ 100 points long in each growing season.

As a final assessment of the methodology, I compared the inferred $K_h$ used to link the two profiles (i.e. $K_h = H/\rho c_p(\partial T_a/\partial z)^{-1}$) to a general model for $K_h$:

$$K_{h,\text{stab}} = \frac{\kappa u^*(z - d)}{\phi(\zeta)}$$  \hspace{1cm} (6.6)

where $\kappa = 0.4$ is the von Karman constant and the stability correction ($\phi(\zeta)$) is formulated after Kader and Yaglom (1990) as a function of the Obukhov length, which may be estimated from measured $H$ and $u^*$. Also included in the comparison is modeled diffusivity for neutral conditions (i.e. $K_{h,\text{neut}} = \kappa u^*(z - d)$). Results are shown in Figure 6.10. When averaged across $u^*$ classes, the inferred diffusivity
Figure 6.10: Eddy diffusivity ($K_h$) as inferred from sensible heat flux measurements and vertical temperature profiles measured with the MMS ($K_{h,MMS}$) and modeled assuming neutral conditions ($K_{h,neut}$) and with a stability correction ($K_{h,stab}$) (top panel). The lower panel shows a histogram of the stability parameter $\zeta$ estimated from $H$ and $u^*$ records from the MMS.

generally matches the near-neutral model better than the stability corrected model, (Figure 6.10a), suggesting that largely unstable or largely stable conditions are not often experienced during the study period. Indeed, a histogram of $\zeta$ (Figure 6.10b) shows that, most of the time, $\text{abs}(\zeta) < 0.2$.

Given that many of the SG flux estimates may be explained by thermodynamic theory, and that the CO$_2$ flux estimates from the SG method are consistent with the $F_{CO2,EC}$ estimates over long time scales, the relative magnitude and direction of the fluxes shown in Figure 6.6 may be robust. However, due to the large uncertainty
associated with the estimates, I caution against placing too much confidence in the absolute value of any one flux estimate produced by either of these two methods on short time scales.

6.6 Conclusions & and an evaluation of methodologies

A perceived trade-off exists between the ecosystems services such as nutrient removal and the emission of GHGs in wetlands ecosystems, and in particular, wetlands constructed as part of compensatory mitigation programs. Here, I employed a range of micro-meteorological techniques to estimate the net flux of trace gases from several hydrologically distinct areas of a restored wetland. I found that water table depth and associated thermodynamic theory for microbial processes explains broad trends in GHG gas fluxes, with more positive fluxes of CO₂, N₂O, and CH₄ observed in inundated areas of the wetland where low redox potentials promote denitrification and methanogenesis, and sparse vegetation limits photosynthetic assimilation of CO₂. I also found that, during the study period, positive effluxes of CH₄ and N₂O from areas of the wetland with deep standing water are largely balanced by sink fluxes in more oxic areas of the ecosystems, and that in all studied areas, CO₂ fluxes were negative or nearly neutral. Noting that at least 50% of the TRP lies above the water table (Figure 6.1), the trade-off of between nutrient services and GHG gases does not appear to be significant in this ecosystem, at least during the peak of the growing season.

These micro-meteorological methods, which can estimate fluxes at scales of 10² - 100² m, were employed largely in an attempt to address uncertainty related to small-scale spatial heterogeneity in wetland GHG flux measurements which infects the more traditional chamber based approach. Since the development of high frequency CO₂ and H₂O gas analyzers, the eddy covariance method has been widely used to measure CO₂ and energy fluxes over a broad range of more than 400 biomes
(Friend et al., 2007), though there is a dearth of EC data from temperate wetlands and boreal peatlands (see the Ameriflux website for a list of North American Eddy Covariance sites http://public.ornl.gov/ameriflux/index.html). In most cases, these are long-term measurement systems that are used to estimated annual fluxes of mass and energy (Friend et al., 2007). Short-term data records may be sufficient to parameterize ecosystem models and estimate fluxes for certain time periods (i.e. Figures 6.4 and 6.5). The methodology is prone to errors, particularly when employed over complex topography with significant vertical and horizontal advection, but these errors are generally understood and should not be a major factor in the relatively flat terrain that characterizes most wetland ecosystems. In any case, the uncertainty associated with the EC method is much less than that associated with the SG method.

Some high-frequency gas analyzers suitable for application with the EC methodology also exist for methane and nitrous oxide, though they are less frequently employed given the relative high cost and significant maintenance demands of these instruments (Lai, 2009; Rinne et al., 2007). Nonetheless, the EC method is a powerful tool to estimate the net ecosystem flux of GHGs from aquatic ecosystems, which hitherto has largely been determined with chamber techniques.

The SG methodology employed here was frequently used to estimated ecosystem fluxes of CO₂ and H₂O before the development of the high frequency gas analyzers that permit EC measurements. It was utilized in this experiment as a low cost alternative to CH₄ and N₂) EC measurements. In this study, the magnitude and direction of fluxes estimated with this methodology are constrained (Figure 6 and 8) with a few exceptions (i.e. the flux of CH4 from TWR₄ in 2008, see Figure 8). The data records are infected with significant uncertainty (on the order of 10 - 20 % for records of ~ 100+ measurements, and much higher uncertainty for shorter records). Thus, future work that relies on the Scalar Gradient technique should be focused on
maximizing the length of the data records.
Conclusions and future work

This dissertation is comprised of five case studies which, while varied in topic and approach, were all conducted with the broad objective of reducing uncertainty in models of the biosphere-atmosphere exchange of greenhouse gases. The key specific findings are summarized in the next section. Following this summary, future research work motivated by the results of these case studies is outlined.

7.1 Summary of results

This work used a combination of field measurements and modeling exercises, with a focus on employing new and emerging techniques, to answer a diverse array of research questions in the field of biophysics. The salient results for the field experiments include a detectable nocturnal evapotranspiration signal in eddy-covariance flux records (Chapter 1), the quantification of the impact of elevated carbon dioxide resin production in field grown pine trees (Chapter 4), and an assessment of the emission rates of the full suite of greenhouse gases (CO₂, H₂O, N₂O, and CH₄) from heterogeneous wetland in Eastern North Carolina (Chapter 5). Modeling work includes an assessment of frameworks for gapfilling nocturnal evapotranspiration...
records (Chapter 1), and the development of prognostic relationships between the parameters of ecosystem carbon and water cycling models and features of canopy architecture and climatic conditions that are readily available at coarse spatial scales (Chapters 3 and 4). In all the case studies, results were constrained and explained with well-established theory, including the cohesion-tension theory for vascular water transport (Chapters 2 and 3), the growth-differentiation balance theory for carbon allocation (Chapter 5), thermodynamic theory describing the energetic requirements of microbial processes (Chapter 5), and micro-meteorological frameworks describing surface layer flow regimes (Chapters 1 and 5). The specific conclusions of each chapter are now summarized.

Chapter 2. Nocturnal evapotranspiration in eddy covariance records.

Until recently, nocturnal water fluxes to the atmosphere were largely ignored in ecosystem water budgets and evapotranspiration models. However, recent evidence suggests that nocturnal transpiration can be a significant component of the total annual evapotranspiration flux (See Figure 1.3). Here, the magnitude and variability of nocturnal evapotranspiration ($ET_N$) is for the first time explored using multiple years of eddy-covariance measurements from three adjacent ecosystems in the Southeastern U.S.: an old grass field, a planted pine forest, and a late-successional hardwood forest. After removing unreliable data points collected during periods of insufficient turbulence, observed $ET_N$ averaged 8 to 9% of mean daytime evapotranspiration ($ET_D$). $ET_N$ was driven primarily by wind speed and vapor pressure deficit and, in the two forested ecosystems, a qualitative analysis suggests a significant contribution arising from nocturnal transpiration rather than soil evaporation. To gapfill missing data, I investigated several methodologies, including process-based multiple non-linear regression, relationships between daytime and nighttime $ET$ fluxes, marginal distribution sampling, and multiple imputation. The utility of the gapfilling proce-
dures was assessed by comparing simulated fluxes to reliably measured fluxes using randomly generated gaps in the data records, and by examining annual sums of ET from the different gapfilling techniques. The choice of gapfilling methodology had a significant impact on estimates of annual ecosystem water use and, in the most extreme cases, altered the annual estimate of ET by over 100 mm yr$^{-1}$, or ca. 15%. While no single gapfilling methodology appeared superior for treating data from all three sites, marginal distribution sampling generally performed well, producing flux estimates with a site average bias error of <10 %, and a mean absolute error close to the random measurement error of the dataset (12.2 and 9.8 W m$^{-2}$, respectively).

Chapter 3. The relationship between reference canopy conductance and simplified hydraulic architecture.

The response of stomatal canopy conductance to fast acting (< 1 hour) meteorological variables is well-studied; however, the processes that govern conductance, and by extension transpiration, on longer time scales are not as well known. Here, combinations of published conductance measurements obtained with several methodologies across boreal to tropical climates were used to explore relationships between canopy conductance rates and hydraulic constraints. A parsimonious hydraulic model requiring sapwood-to-leaf area ratio and canopy height generated acceptable agreement with measurements across a range of biomes ($r^2 = 0.75$). The results suggest that, at long time scales, the functional convergence among ecosystems in the relationship between water-use and hydraulic architecture eclipses inter-specific variation in physiology and anatomy of the transport system, and features of hydraulic architecture may be used to improve $PFT$-based approaches to the estimation of model parameters describing reference canopy conductance rates. Prognostic applicability of this model requires independent knowledge of sapwood-to-leaf area. In this study, I did not find a strong relationship between sapwood-to-leaf area and physical or climatic
variables that are readily determinable at coarse scales, though the results suggest that climate may have a mediating influence on the relationship between sapwood-to-leaf area and height. Within temperate forests, canopy height alone explained a large amount of the variance in reference canopy conductance ($r^2 = 0.68$) and this relationship may be more immediately applicable in the terrestrial ecosystem models.

Chapter 4. A characterization of the variability in ecosystem-scale model parameters using Bayesian inversion of eddy covariance data.

Parameterization schemes remain one major source of uncertainty in the output of DGVMs. In this chapter, I applied Bayesian inversion techniques to eddy covariance records from over 35 diverse ecosystems to determine the parameters of simple models for the net ecosystem exchange of CO$_2$ (NEE) and canopy stomatal conductance ($G_T$). The parameters varied significantly across sites, with an average coefficient of variation of 0.37. Grouping parameters on the bases of PFT did little to reduce this variation, as intra-PFT coefficients of variation remain large (C.V. $\sim 0.30$) and ranges for the parameters overlap considerably across PFTs. Significant inter-annual variation in these model parameters was also observed. I also explored relationships between the model parameters and climatic and structural indices that are readily available at coarse-spatial scales in an effort to identify prognostic relationships that could reduce parameter uncertainty. Reference canopy conductance ($G_{Trref}$), maximum assimilation rate ($\beta$), and ecosystem respiration rates ($\gamma$) were all significantly related to canopy height ($r^2 = 0.46$, 0.39, and 0.34, respectively), a result that can be explained using cohesion-tension theory and coordination between the photosynthetic and hydraulic apparatus of the plant. Hourly estimates of the flux-based marginal water use efficiency were significantly related to the inverse square root of vapor pressure deficit within many sites, and across sites, marginal water use efficiency was explained in part by long-term averaged vapor pressure deficit ($r^2 = 0.27$),
a result that suggests a functional convergence in marginal water use efficiency across the study sites when analyzed within a recently developed theoretical framework for optimal stomatal functioning. I concluded by showing that these inter-specific relationships can improve the distribution of errors between modeled and measured flux rates when compared to a purely PFT-based approach to parameterization.

Chapter 5. Elevated atmospheric CO$_2$ enhances resin production in field-grown pine trees.

Bark beetles are among the most significant disturbance agents in forest ecosystems, and warmer temperatures associated with future climates are expected to promote more beetle outbreaks. However, climate-induced changes in resin production, by which pine trees defend against bark beetle attack, are not well known. In this study, the Free-Air Carbon Enrichment (FACE) facility was used to assess the impact of elevated atmospheric CO$_2$ and nitrogen amendments on resin production in a southeastern U.S. loblolly pine plantation. Resin flow in canopy dominant trees exposed to elevated CO$_2$ was significantly higher than resin flow in trees grown in ambient CO$_2$, with an enhancement >100 % in unfertilized trees. Resin production also varied significantly with time of year (P < .01) and with tree size (P < .001), and was strongly correlated to aboveground woody net primary productivity ($r^2 = 0.64$ to 0.90 for canopy dominants) during the growing season. These trends can be explained by the growth-differentiation balance hypothesis - a theoretical framework that describes resource allocation to defense mechanisms and that is generally applicable to other plant-herbivore systems.

Chapter 6. Effects of water table depth variations on ecosystem-scale fluxes of greenhouse gases in a large restored temperate wetland.

The effluxes of greenhouse gases like CH$_4$ and N$_2$O are not well constrained (Table 1.1). In this chapter, two atmospheric flux-estimation techniques that can inte-
grate over smaller scale spatial heterogeneities (the eddy covariance methodology and a variant of the flux-gradient technique) were employed to determine fluxes of CO$_2$, H$_2$O, CH$_4$, and N$_2$O over hydrologically distinct areas within a large restored wetland in the Southeastern U.S. In general, the magnitude and direction of flux estimates followed predictions from thermodynamic theory. Sinks for CO$_2$ observed for all study areas, with mean study period magnitudes ranging from -88 to -407 mg CO$_2$-C m$^{-2}$ h$^{-1}$ were observed for all study areas. Significant sink fluxes of CH$_4$ (with a study period average of -7.7 ± 3.5 to -4.2 ± 2.5 mg CH$_4$-C m$^{-2}$ h$^{-1}$ for CH$_4$) were observed in a dryland area of the wetland and source fluxes of CH$_4$ (with a study period average of 1.33 ± 3.5 to 5.2 ± 2.2 mg CH$_4$-C m$^{-2}$ h$^{-1}$) were observed in an inundated portion of the wetland. N$_2$O fluxes were also estimated to be a sink or source depending on water table depth. Evapotranspiration estimates were similar for a dryland and inundated area of the wetland, with evidence that increased transpiration in the dryland area may compensate for increased evaporation in the flooded area. Due largely to a balance of sink and source fluxes, I concluded that this wetland is not a significant source of greenhouse gases during the growing season. The error associated with a micro-meteorological flux-gradient method employed here was large. Nonetheless, when averaged over longer time scales, these atmospheric estimates of wetland GHG fluxes provide an important constraint on highly variable chamber-based flux estimates provided a sufficient number of data have been collected.

7.2 Future work

The results outlined in the previous section, together with the broad research needs outlined in the introduction, motivate future research objectives, some of which I intend to address during my tenure as a post-doctoral researcher with the U.S. Forest Service. These include:
The interpretation of nocturnal eddy covariance records, and the application of eddy covariance methodologies over complex terrain. Nocturnal eddy covariance measurements must routinely be rejected due to the invalidation of the simplifying assumptions to the mean scalar continuity equation outlined in Section 6.2.1. However, as the results from Chapter 1 illustrate, important yet poorly understood processes that occur at night may be important contributors to annual flux estimates. The difficulties in interpreting nocturnal eddy covariance data are increased when flux towers are situated in complex topographies, where drainage and advective flows may influence not only the interpretation of flux measurements, but also the physiological processes that drive them (Aubinet, 2008). For this reason, most eddy covariance flux towers have been preferentially located in homogeneous ecosystems with little topographical variation (Baldocchi, 2008). The challenge facing the eddy covariance community is to expand the network to include more heterogeneous and complex topographies, and develop the operational standards and theoretical frameworks necessary to properly interpret these datasets (see Katul et al. 2006 for a relevant discussion). As part of my work with the U.S. Forest Service, I plan to oversee such an effort by assisting in the development of an eddy covariance tower and ancillary gas profile systems in exceptionally complex terrain within the Coweeta Long-Term Ecological Research facility in mountainous western North Carolina. I am currently exploring what has been learned from the advection studies conducted across Euroflux sites and assessing whether an operational data assimilation framework that combines simplified theories for advection and multiple measurement locations may be a logical first step (vis-a-vis measuring the advective terms explicitly to correct for the canopy sources and sinks).
• The application of eddy covariance methodologies to understudied GHGs. Quantification of trace gases like CH$_4$ and N$_2$O from wetland ecosystems is an important research need (Denman et al., 2007). In Chapter 5, a variant of the flux-gradient methodology is employed to estimate the fluxes of CH$_4$ and N$_2$O from a large temperature wetland in Eastern North Carolina. The flux-gradient techniques was historically used to estimate the ecosystem exchange of CO$_2$ before the advent of high frequency CO$_2$ gas analyzers, and was employed in this study as a low-cost alternative to eddy covariance systems for CH$_4$ and N$_2$O, which have only recently been developed and which suffer from a number of operational difficulties (Rinne et al., 2007). However, as our results show, the uncertainty associated with this technique is quite large when compared to the eddy covariance measurements, and the development of networks of eddy covariance systems for measuring these other trace gases should become a research priority. In some areas, this effort is already underway (Soussana et al., 2007), and I will dedicate some effort to be part of this network in North America.

• Improved parameterization schemes for DGVMs. As discussed in Chapters 1, 3, and 4, a major source of error in the current suite of DGVMs is the improper specification of model parameters. As I show in Chapters 3 and 4, model parameters vary considerably within species, within PFTs, and with time. But in some cases theoretically-based relationships exist between model parameters and information that is readily available at coarse spatial scales that may better constrain parameter specification in regional- to global-scale applications. Specifically, canopy height, which is available as regional scale datasets for most of the terrestrial land surface (Kellndorfer et al., 2004), appears capable of providing a necessary constraint on many carbon and water cycling param-
eters, and including these relationships in coarse-scale simulations in a natural next step. Further, other workers have had similar successes (Oren et al., 1999b; Katul et al., 2009; Meinzer et al., 2005) in elucidating inter-specific convergences for key model parameters. Given the large number of parameters required by DGVMs (Table 1.2), rich opportunities exist for additional efforts to develop theoretically-based improvements to parameterization schemes.

• *Characterizing the risk of insect outbreaks in current and future climates* In Chapter 5, for the first time, the impacts of changing atmospheric CO$_2$ concentrations on resin flow in pine trees was quantified. A natural next step is to incorporate these results into models that predict the risk or damage from bark beetle outbreaks in future climates. Additionally, results from Chapter 5 show that resin flow varies on short time scales (i.e. monthly) and is negatively correlated with growth processes, which in turn are driven by meteorological and edaphic conditions. Hence, these results may also inform efforts to assess levels of susceptibility and risk to insect outbreaks in the present. Towards that end, as part of my work with the Forest Service, I will be working to develop metrics to assess the degree of forest stress from a unique network of sap flux and meteorological measurements that are updated in real time via satellite links. Finally, I note that forest-pest dynamics is a rich, challenging, and complex field of research with many unanswered questions and opportunities. Some areas I find particularly interesting include spatial models for the spread of insect outbreaks, and interactions between pests, host trees, and other organisms like insect predators or the symbiotic fungi that are linked to bark beetle infestations.
Appendix A

Sensitivity Analysis for Reference Conductance

To assess the sensitivity of $G_{Sref}$ to $A_S/A_L$, $\Psi_{leaf}$, $k_s$, and $h$, consider a Taylor series expansion of $G_{Sref}$:

$$\partial G_{Sref} = \frac{\partial G_{Sref}}{\partial A_S/A_L} \delta A_S/A_L + \frac{\partial G_{Sref}}{\partial \Psi_{leaf}} \delta \Psi_{leaf} + \frac{\partial G_{Sref}}{\partial k_s} \delta k_s + \frac{\partial G_{Sref}}{\partial h} \delta h.$$

(A.1)

Upon computing all the partial derivatives in Equation A.1 using Equation 3.4 and expressing the outcome as relative changes, the above equation simplifies to

$$\frac{\delta G_{Sref}}{G_{Sref}} = \frac{\delta A_S/A_L}{A_S/A_L} + \frac{\delta k_s}{k_s} + \frac{1}{1 - \rho_w gh(\Psi_{leaf})^{-1}} \left( \frac{\delta \Psi_{leaf}}{\Psi_{leaf}} - \frac{\delta h}{h} \right).$$

(A.2)

Equation A.2 analytically demonstrates that the relative change in $G_{Sref}$ scales linearly with the relative changes in $A_S/A_L$ and $k_s$, but not with $\Psi_{leaf}$ and $h$. Using typical literature values as 'reference states' ($\Psi_{leaf} = -2$ MPa, $k_s = 3$ m$^2$, $h = 20$ m and $A_S/A_L = 4$ cm$^2$ m$^{-2}$), Equation A.2 is evaluated for a range of values bounded by the extremes cited in the text. The results suggest that $G_{Sref}$ varies by a factor...
of \sim 10 with \( h \), by a factor of \sim 3.5 with \( A_S/A_L \), and by a factor of \sim 0.5 with \( \Psi_{leaf} \) and \( k_s \). Stated differently, the sensitivity analysis in Equation A.2 demonstrates that when considering the reported variations in the literature in each of these parameters across species, \( \delta k_s/k_s < < \frac{\delta A_S/A_L}{A_S/A_L} \) and \( \delta \Psi_{leaf}/\Psi_{leaf} < < \delta h/h \), although this argument need not hold for all species.

Nevertheless, among many species a reasonable approximation is:

\[
G_{Sref} \approx \frac{\delta A_S}{A_L} - \frac{1}{1 - \rho_w gh (\Psi_{leaf}^{-1})} \frac{\delta h}{h}.
\]  

(A.3)

As expected, Equation A.3 analytically predicts that \( G_{Sref} \) diminishes rapidly with increasing height for small \( h \) if no adjustments in \( A_S/A_L \) occur. The quantity \( 1/(1 - \rho_w gh (\Psi_{leaf}^{-1})^{-1}) \) can be assumed to equal unity for a wide range of ecosystems, noting that \( \rho_w \approx 10^3 \text{ kg m}^{-3} \), \( g \approx 10 \text{ m s}^{-2} \), and \( \Psi_{leaf} \approx 10^6 \text{ kg m}^{-2} \text{ s} \). This approximation can be used to explicitly assess the relative contribution of \( \partial h/h \) and \( \partial A_S/A_L/A_S/A_L \) to \( \partial G_{Sref}/G_{Sref} \) within a species.
Appendix B

The drivers of resin flow elucidated from a hierarchical Bayesian analysis

We used a hierarchical Bayesian framework in addition to the ANOVA analysis to determine the impacts of elevated CO\textsubscript{2}, fertilization, the interaction between atmospheric CO\textsubscript{2} and fertilization, and sampling date on mean resin production. Within this framework, we also investigated the contribution of individual tree effects versus random error to the total error in the dataset.

The model is a generalized linear model with fixed effects for treatment co-factors, fixed month effects ($\kappa_{R,t}$), random individual effects ($\beta_{R,i}$), and random error ($\varepsilon_{R,it}$):

$$MR_{i,t} = (m_R)_{i,t}^{1/2} = X_{i,t}n_R + \beta_{R,i} + \kappa_{R,t} + \varepsilon_{R,it}$$ (B.1)

where $i = 1,...,n$ (with $n = 160$ is the number of resin samples each month), and $t = 1,...,T$ (where $T = 8$ is the number of sampling months). The dataset is transformed by the square root function to ensure the marginal distributions of the data within the treatment classes are approximately Gaussian. $X_{R,it}$ is the vector of fixed effects for sample $i$ in month $t$ ($X_{it} = [S_{it} C_{it} N_{it} (C \times N)_{it}]$ ) where $S_{i,t}$ represents...
size class (1 = dominant, 0 = suppressed), $C_{i,t}$ and $N_{i,t}$ represent binary CO$_2$ and fertilization effects, respectively, and $C_{i,t} \times N_{i,t}$ represents their interaction. $\alpha$ is the corresponding vector of regression parameters ($\alpha_R = [\alpha_{R,S} \ \alpha_{R,C} \ \alpha_{R,N} \ \alpha_{R,C \times N}]$). Both the individual effects and random error are assumed to be normally distributed with mean zero and variance parameters $\tau^2_R$ and $\sigma^2_R$, respectively, yielding:

\[ \beta_i \sim N(\beta_i | 0 \ \tau^2_R ) \] (B.2)

\[ \varepsilon_{it} \sim N(\varepsilon_{it} | 0 \ \sigma^2_R ) \] (B.3)

The full probability model employed here is very similar to the "conditional fecundity" model employed by Ladeau and Clark (2006) and we refer readers to that publication for additional computation details. Briefly, conjugate, non-informative priors were employed for all of the model parameters. The regression parameters were drawn from a very flat multivariate normal distribution with mean vector $\alpha_\theta = [0 \ 0 \ 0 \ 0]$ and covariance matrix $V_\alpha = \text{Diag}[1000 \ 1000 \ 1000 \ 1000]$. Priors for month effects, tree effects, and random effects were all normally distributed with zero mean. For $\kappa_R$, the prior variance was set to $v_K = 1000$. For $\beta_{R,i}$ and $\varepsilon_{R,it}$, the variance was modeled with inverse gamma hyperpriors for the variance:

\[ \tau^2_R = IG(\tau^2_R | .1 \ .1 ) \] (B.4)

\[ \sigma^2_R = IG(\sigma^2_R | .1 \ .1 ) \] (B.5)

The conditional posteriors for the model parameters are all normally distributed with mean and variance parameters as described in Clark (2007) and Ladeau and Clark (2006). The posteriors were sampled using a Gibbs Sampler with 120,000
iterations. The parameter chains were then thinned by extracting every 200th sample. The first half of the thinned chains were then discarded, leaving 300 parameter estimates to build the posterior PDFs.

The inversion was performed twice: first using all available resin samples, and then using resin samples from dominant trees alone. In the latter case, $X_{it} = [C_{it} \ N_{it} \ C \times N_{it}]$. Histograms of the regression parameters for both cases are shown in Figure B1. When all data are considered, the class size effect is positive ($\langle \alpha_{R,S} \rangle = 0.23$) with a 95% confidence interval that ranges from 0.01 to 0.44. For both simulations, the CO$_2$ enrichment effect is also positive ($\langle \alpha_{R,C} \rangle = 0.31$ and 0.45 for all data and dominants only, respectively), though we note that the 95% confidence interval lies entirely above zero only when large trees alone are considered (Figure B.1e). For both datasets, estimates of the fertilization parameter is generally positive and the C x N interaction term is generally negative. However, the confidence interval for the fertilization effect parameter includes both positive and negative estimates for both simulations.

The posterior distributions for $\sigma^2_R$ and $\tau^2_R$ are similar for both distributions (Figure B.2). Average values of these parameters are $\langle \sigma^2_R \rangle = 0.16$ and 0.17, and $\langle \tau^2_R \rangle = 0.11$ and 0.11 for the simulations will all data and with canopy dominants, respectively. Hence, individual tree variation explains roughly 40% of the error not explained by the treatment co-factors and sampling date, with the remainder attributable to spatial variation in resin flow within a tree, and random measurement error.

The month effects $\kappa_R$ are shown in Figure B.3 for the first simulation, noting that the results from the second are nearly indistinguishable. The confidence intervals on these parameters are quite large and reflect a large degree of correlation among distributions of these monthly parameters ($r^2 = 0.72$ to 0.80). However, the trend in these time effects is quite similar to the time series produced by a simple averaging of $m_R$ (i.e. Figure 5.1).
Figure B.1: The posterior distributions of the regression parameters for size effects (S), CO$_2$ effects (C), fertilization effects (N), and the CO$_2$ x fertilization effect (C x N). The top panel (a-d) shows results from the model that includes a size effect; the lower panel (e-g) shows results from the model that includes dominant trees only.
Figure B.2: The posterior distributions of the variance parameters for random individual effects ($\tau^2_R$) and random error ($\sigma^2_R$). The top panel (a-b) shows results from the model that includes a size effect; the lower panel (c-d) shows results from the model that includes dominant trees only.
Figure B.3: The month effects from the model that includes a size effect. Symbols show the mean and error bars show the 95% confidence intervals on the mean.

Finally, we note that the tree effects $\beta_{R,i}$ are strongly correlated with the mean $m_R$ for each tree averaged over months ($r^2 \sim 0.80$, data not shown), illustrating the inference capabilities of the Bayesian hierarchical approach even when employed with a dataset characterized by a large degree of random error.
Appendix C

Simulation exercise to explore SG errors associated with concentration measurement errors

To determine how concentration measurement errors affect flux estimates using the SG methodology, scalar concentrations were artificially generated for heights of 2 m, 3 m, and 5 m from normal distributions assuming a log-linear relationship between scalar concentration and height based on observed profiles for methane over TWR\textsubscript{A} (i.e. the 'true' dataset). Next, three additional datasets were simulated and injected with 1%, 2%, and 4% error, which spans the range in observed concentration errors (Table 6.3) from the concentration profiles in the TRP wetland.

Next, I simulated eddy diffusivities on an hourly time scale using meteorological data recorded from the MMS in 2009. Specifically, I assumed the widely used formulation for $K_h$ in near-neutral conditions (Miyata et al., 2000; Stull, 1988):

$$K_h = \kappa u^*(z - d)$$  \hspace{1cm} (C.1)

where $\kappa=0.4$ is the von Karman constant.

The scalar flux for all four datasets was then estimated via Equation 6.2, and
the absolute error between the mean 'true' flux and the simulated datasets was determined for data records of increasing length, beginning at 20 measurements and ending at 200. This procedure was repeated using a Monte Carlo analysis with 1000 simulations. The results are shown in Figure C.1 and suggest that, on average, at least 100 data points are needed to recover the 'true flux' mean to within 10% if the concentration error is 1%. Relative errors near 20% persist even for data records that are over 100 points long for 2% and 4% concentration errors.
Bibliography


M. Alsheimer, B. Köstner, E. Falge, and J. D. Tenhunen. Temporal and spatial variation in transpiration of Norway spruce stands within a forested catchment of


P. D. Blanken and T. A. Black. The canopy conductance of a boreal aspen forest,


G. J. Collatz, J. T. Ball, C. Grivet, and J. A. Berry. Physiological and environmental-regulation of stomatal conductance, photosynthesis and transpiration - a model that includes a laminar boundary-layer. *Agricultural and Forest Meteorology*, 54


A. J. Elmore, J. F. Mustard, S. J. Manning, and D. B. Lobell. Quantifying vegeta-


L. H. Gu, T. Meyers, S. G. Pallardy, P. J. Hanson, B. Yang, M. Heuer, K. P. Hosman, J. S. Riggs, D. Sluss, and S. D. Wullschleger. Direct and indirect effects


K. D. Klepzig, D. J. Robison, G. Fowler, P. R. Minchin, F. P. Hain, and H. L. Allen.


T. Kumagai, M. Tateishi, T. Shimizu, and K. Otsuki. Transpiration and canopy con-


X. H. Lee and X. Z. Hu. Forest-air fluxes of carbon, water and energy over non-flat


M. J. Lombardero, M. P. Ayres, B. D. Ayres, and J. D. Reeve. Cold tolerance of four


D. T. Mihailovic, B. Rajkovic, B. Lalic, and L. Dekic. Schemes for parameterizing


M. Reichstein, E. Falge, D. Baldocchi, D. Papale, M. Aubinet, P. Berbigier, C. Bern-


M. G. Ryan, B. J. Bond, B. E. Law, R. M. Hubbard, D. Woodruff, E. Cienciala, and J. Kucera. Transpiration and whole-tree conductance in ponderosa pine trees of


M. B. S. Siqueira, G. Katul, and A. Porporato. Soil moisture feedbacks on convective


S. C. Thomas and W. E. Winner. Photosynthetic differences between saplings and


M. T. Tyree. The cohesion-tension theory of sap ascent: current controversies. *Jour-


Y. P. Wang, R. Leuning, H. A. Cleugh, and P. A. Coppin. Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful? *Global Change Biology*, 7(5):


density effects due to heat and water-vapor transfer. *Quarterly Journal of the

L. A. Wever, L. B. Flanagan, and P. J. Carlson. Seasonal and interannual variation
in evapotranspiration, energy balance and surface conductance in a northern

S. C. Whalen. Biogeochemistry of methane exchange between natural wetlands and


Xu-Ri and I. C. Prentice. Terrestrial nitrogen cycle simulation with a dynamic global


Kimberly A. Novick was born in July, 1980 in Montgomery County, Maryland and graduated from Quince Orchard High School in Gaithersburg, Maryland in 1998. She received a Bachelor of Science in Engineering degree from Duke University in 2002. She currently lives in Durham, N.C. with her partner and son. In the summer of 2010, she will begin a post-doctoral position with the U.S. Forest Service.

When she wasn’t running MatLAB scripts as part of her dissertation research, she enjoyed the opportunity to serve in a number of university- and community-oriented volunteer positions. These activities included: serving as the Nicholas School representative to the Graduate and Professional Student Council (GPSC) for two years, serving on the GPSC Young Trustee Selection Committee, the GPSC Community Outreach Committee, and volunteering with the GPSC Student Health Advisory Committee. During her final year, she also served as a graduate student representative to the Duke Board of Trustee’s subcommittee on Academic Affairs. Within the Nicholas School, she chaired the graduate student Orientation and Social Committees.

Outside of academia, she serves on the organizing committee for the Troika Music Festival. In 2008, she volunteered hundreds of hours to register and turn out voters with the Obama campaign. She also enjoys sewing handbags and hosting a weekly trivia quiz in her local Irish Pub.