The Pennsylvania State University
The Graduate School
Department of Meteorology

DIAGNOSING UNCERTAINTY AND IMPROVING PREDICTIONS OF TERRESTRIAL CO₂ FLUXES AT MULTIPLE SCALES THROUGH DATA ASSIMILATION

A Thesis in
Meteorology
by
Daniel Michael Ricciuto

© 2006 Daniel Michael Ricciuto

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

August 2006
The thesis of Daniel Michael Ricciuto was reviewed and approved* by the following:

Kenneth J. Davis  
Associate Professor of Meteorology  
Thesis Advisor  
Chair of Committee

Raymond G. Najjar  
Associate Professor of Meteorology and Geosciences

John C. Wyngaard  
Professor of Meteorology, Mechanical Engineering and Geoenvironmental Engineering

Klaus Keller  
Assistant Professor of Geosciences

William H. Brune  
Professor of Meteorology  
Head of the Department of Meteorology

*Signatures are on file in the Graduate School
ABSTRACT

We use data assimilation techniques to assimilate carbon cycle data over a range of spatial and temporal scales in three models. First, global CO₂ concentration data and ocean flux observations are used to calibrate a simple carbon cycle model, to estimate uncertainty in key carbon cycle parameters, and to predict the behavior of both the terrestrial and ocean carbon sinks. Given equal observation errors, historical temperature forcing, and our assumed model structure, hypothetical annual observations of global terrestrial CO₂ fluxes reduce predictive uncertainties about CO₂ sinks more than twice as much as annual observations of global ocean CO₂ fluxes. The main reason for this effect lies in the interannual variability of the terrestrial carbon cycle that is resolved by our model and that constrains model parameters. This conclusion hinges on the model structure that assumes much smaller interannual variability of the oceanic carbon cycle.

Direct observations of terrestrial CO₂ fluxes do not exist at a global scale, and data assimilation experiments using direct terrestrial flux observations must focus on smaller scales until appropriate upscaling techniques are discovered. Eddy covariance towers measure continuous fluxes of CO₂ with a spatial footprint on the order of 1 km². At the WLEF eddy covariance tower, which is in a northern Wisconsin mixed forest, interannual variations in observed sums of CO₂ flux are found to be statistically significant with respect to random error. These flux data were assimilated into a simple climate-driven model. Considerable differences in respiration parameter probability density functions (PDFs) occurred depending on whether daytime, nighttime, or all flux data were used in the assimilation. This simple model was moderately successful in
producing statistically significant correlations with interannual variations in annual and

growing season net ecosystem exchange (NEE) sums, but was generally unsuccessful in

spring and autumn. In all cases, the model underestimated the range of variability in

NEE sums.

Upscaling eddy covariance tower fluxes from the scale of an eddy covariance

footprint to a regional or global scale is a nontrivial task. Assimilating eddy covariance

data into a process-based terrestrial carbon cycle model is one way to attempt this type of

upscaling. We use the Top-down Representation of Interactive Foliage and Flora

Including Dynamics (TRIFFID) to assimilate eddy covariance data from five sites in

eastern North America. Because of the complexity of the model, only a subset of model

parameters could be optimized; 22 parameters assumed to contribute the most to

observed flux variance are selected. The system is nonconvex, displaying multiple

maxima in the likelihood function. We use a genetic algorithm to find the global

optimum set of parameters for each site individually, and jointly among all five sites.

Variability from diurnal to seasonal timescales is simulated with high accuracy by the

optimized models, a large improvement over the results obtained when using published

parameters. There was little skill in reproducing interannual variability using either the

published or the optimized parameters, indicating that key processes governing

variability at this timescale are, as of yet, missing in TRIFFID.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... vii

LIST OF TABLES ............................................................................................................ xii

PREFACE ...................................................................................................................... xiv

ACKNOWLEDGEMENTS ............................................................................................... xv

Chapter 1  Introduction ................................................................................................ 1

  1.1 Background ..................................................................................................... 1
  1.2 Outline ........................................................................................................... 5

Chapter 2  A Bayesian synthesis inversion of carbon cycle observations: How can observations reduce uncertainties about future sinks? ........................................... 7

  2.1 Introduction ..................................................................................................... 7
  2.2 Methods and Data .......................................................................................... 12
      2.2.1 Carbon cycle model ............................................................................. 12
      2.2.2 Model forcings .................................................................................... 17
      2.2.3 Observational Constraints .................................................................. 18
      2.2.4 Data assimilation technique ............................................................... 19
      2.2.5 Hypotheses and experimental setup .................................................... 24
  2.3 Results and Discussion ................................................................................... 27
      2.3.1 Base case parameter distributions ....................................................... 27
      2.3.2 Effects of neglecting autocorrelation .................................................... 33
      2.3.3 Effects of interannually varying temperature ....................................... 35
      2.3.4 Model performance ............................................................................ 36
      2.3.5 Probabilistic predictions ..................................................................... 39
  2.4 Conclusion ....................................................................................................... 43

Chapter 3  Causes of interannual variability in ecosystem-atmosphere carbon dioxide exchange in a northern Wisconsin forest using a Bayesian synthesis inversion ...................................................... 45

  3.1 Introduction ..................................................................................................... 45
  3.2 Methods .......................................................................................................... 49
      3.2.1 Data sources .......................................................................................... 49
      3.2.2 Error analysis ....................................................................................... 50
          3.2.2.1 Data screening ............................................................................. 51
          3.2.2.2 Gap-filling methodology ............................................................. 54
          3.2.2.3 Gap-filling uncertainty ................................................................. 56
          3.2.2.4 Turbulent sampling uncertainty ............................................... 57
3.2.2.5 Level selection uncertainty ............................................................... 59
3.2.2.6 Total random uncertainty ................................................................. 63
3.2.3 Interannual model ............................................................................... 63
  3.2.3.1 Respiration submodel .................................................................... 64
  3.2.3.2 Photosynthesis submodel .............................................................. 68
3.2.4 Nonlinear optimization methodology ................................................. 70
3.2.5 Selecting Bayesian Priors ................................................................. 72
3.3 Results and discussion ........................................................................... 74
  3.3.1 Climate .............................................................................................. 74
  3.3.2 Seasonal and interannual variability of gap-filled NEE ...................... 76
  3.3.3 Driving factors of interannual variability ........................................... 79
  3.3.4 Modeling results .............................................................................. 81
  3.3.5 Conclusion ....................................................................................... 90

Chapter 4 Improving model estimates of spatial and temporal variations in CO₂ fluxes: A multi-site data assimilation of TRIFFID .................................................. 93

  4.1 Introduction .......................................................................................... 93
  4.2 Materials and Methods ....................................................................... 98
    4.2.1 Site descriptions ............................................................................ 98
    4.2.2 TRIFFID model description .......................................................... 102
      4.2.2.1 Photosynthesis ..................................................................... 105
      4.2.2.2 Respiration ......................................................................... 107
      4.2.2.3 Phenology .......................................................................... 109
    4.2.3 Model forcings ............................................................................. 110
    4.2.4 Model initialization ...................................................................... 111
    4.2.5 Observational Constraints ........................................................... 112
    4.2.6 Optimization Method ................................................................. 112
    4.2.7 Experimental design ................................................................. 115
  4.3 Results and discussion ....................................................................... 117
    4.3.1 Nonconvexity of likelihood function .......................................... 117
    4.3.2 Convergence of SRES algorithm ................................................. 117
    4.3.3 Parameter estimates .................................................................... 122
    4.3.4 Variability from diurnal to seasonal timescales ......................... 128
    4.3.5 Interannual variability ................................................................. 131
    4.3.6 Intersite variability ..................................................................... 135
  4.4 Conclusion .......................................................................................... 135

Chapter 5 Conclusion .................................................................................... 138

  5.1 Summary ............................................................................................ 138
Bibliography .................................................................................................. 142
LIST OF FIGURES

Figure 2-1: Model schematic diagram showing the climate and carbon cycle modules, where the carbon cycle module is subdivided into terrestrial and ocean components. Constraints are illustrated by shaded ovals, while emissions forcing are white ovals. Observations are used to constrain four model parameters. .............................................................................................................13

Figure 2-2: Schematic of the terrestrial carbon cycle model. NPP is a function of atmospheric CO₂ concentration, and is allocated to leaf and wood pools. These pools turn over at specified rates into the detritus pool and soil carbon pools. These pools then turn over in the atmosphere at a rate dependent on the global mean temperature. Preindustrial pool sizes are given in bold. ...............15

Figure 2-3: Probability density functions (PDFs) of (a) respiration temperature sensitivity Q₁₀, (b) carbon fertilization parameter β, (c) thermocline transfer velocity η and (d) autocorrelation parameter θ. Shown are results for three cases: with autocorrelation and driven by (i) historical temperature (solid), (ii) without autocorrelation and driven by historical temperature (dashed), and (iii) with autocorrelation but driven by modeled temperature (dash-dotted). Comparing cases (i) and (ii) indicates that ignoring autocorrelation in the Mauna Loa CO₂ results in highly biased results. The sharper PDFs in case (i) versus case (iii) imply that the interannual variability in the observed temperature record is a useful source of information about carbon cycle parameters. .............................................................................................................28

Figure 2-4: Scatter plots representing the joint 95% confidence bounds of (a) respiration sensitivity Q₁₀ and carbon fertilization factor β, (b) Q₁₀ and thermocline diffusivity η, (c) β and η, (d) η and the autocorrelation parameter θ, (e) β and θ, and (f) η and θ. ..................................................................................................................32

Figure 2-5: Residuals (the difference between observed and most likely model values) of Mauna Loa CO₂ concentration observations from 1960-2004 (a). The partial autocorrelation function of these residuals is shown in panel (b). The likelihood function used for model calibration has been modified to estimate the lag 1 autocorrelation coefficient θ, (c.f. Figure 1, d). Panels (c) and (d) depict the residuals and the associated partial autocorrelation function after whitening using the maximum likelihood estimate of θ, respectively. Asterisks in panels (b) and (d) delineate the 95% confidence limits based on the null hypothesis of zero partial autocorrelation................................................34

Figure 2-6: Maximum likelihood estimate (solid line) and 95% confidence intervals (dashed lines) of model predictions with observational constraints (asterisks with error bars) for (a) CO₂ concentrations with observations from
Law Dome and Mauna Loa, (b) net terrestrial carbon uptake, (c) ocean carbon uptake with the estimates of McNeil et al. (2003) and (d) cumulative ocean with the estimate of Sabine et al. (2004). The model was driven with historical emissions from 1850-2004.

Figure 2-7: Predicted allowable emissions under the S550 stabilization scenario (a). The maximum likelihood (solid line) and 95% confidence intervals (dashed lines) of model predictions under an S550 stabilization scenario from the years 2005-2100 for ocean + terrestrial carbon uptake (b), terrestrial carbon uptake (c) and ocean carbon uptake (d). In these model simulations, CO₂ concentrations are prescribed and the allowable CO₂ emissions are calculated from the atmospheric balance of CO₂.

Figure 2-8: 95% confidence range of the allowable CO₂ emissions for the S550 stabilization scenario from the years 2005-2100. The solid black line is an estimate based on the currently available historical constraints of CO₂ concentrations and oceanic uptake. The dashed and dotted lines are estimates with the historical constraints plus an added hypothetical 2004 terrestrial flux constraint or an added hypothetical 2004 ocean flux constraint, respectively. Uncertainties of the hypothetical terrestrial and ocean flux constraints are both 0.4 PgC, which is equal to the McNeil et al. constraint uncertainty.

Figure 3-1: Cumulative annual NEE (a) and percentage of hours gap-filled (b) as functions of the u* threshold that is used to screen nighttime data. The dashed vertical line indicates the value of the u* threshold that was used in this paper: 0.2 m s⁻¹.

Figure 3-2: Diurnal cycle of estimated hourly random error (turbulent sampling plus instrumental noise) in NEE measurements in December-March (a), April-May (b), June-August (c) and September-November (d). Hourly random error was estimated as the standard deviation of the difference between modeled and observed NEE over each season. Modeled values of NEE were calculated using gap-filling regression parameters.

Figure 3-3: Diurnal cycle of NEE as measured at 30, 122 and 396 meters during winter (a), spring (b), summer (c), and autumn (d). Data are from hours during which measurements exist for all three levels from 1997 through 2004.

Figure 3-4: Bin-averaged dependence of daily R₁₀ and GEPₘₐₓ parameters on soil temperature (Tₛ) and soil moisture (SWC) when assuming the air temperature sensitivity Eₜ and half saturation constant PAR₁/₂ are constant. The Tₛ dependence of R₁₀ is described well by a linear fit represented by the solid line (a). The residuals of this fit are a function of SWC (b). Similarly, the Tₛ dependence of GEPₘₐₓ is described well by a linear fit represented by the solid line (c), and the residuals are a function of SWC (d).
Figure 3-5: Probability density functions (PDFs) of interannual model parameters estimated by the MCMC data assimilation technique. The results of three assimilation experiments are shown: the base-case assimilation with all data (thick solid), with nighttime data only (dashed), and with daytime data only (gray). .................................................................82

Figure 3-6: Dependence of $R_{10}$ on soil temperature as estimated by the interannual model (a) for the base-case assimilation (thick solid), nighttime-only assimilation, and (dashed) daytime-only assimilation (gray). Dependence of GEP at a value of $\text{PAR} = 1000 \, \mu\text{mol m}^{-2} \, \text{s}^{-1}$ ........................................85

Figure 3-7: Modeled and observed daily-averaged NEE values for September 1997 using the interannual model calibrated using the base-case assimilation (a), and modeled and observed monthly averaged-NEE values over all years using the same model (b). Observed values are gap-filled using standard methods when necessary. .......................................................................................87

Figure 3-8: Modeled vs. observed annual NEE (a), annual nighttime NEE (b), and annual daytime NEE (c). The interannual model systematically underestimates the magnitudes of daytime and nighttime fluxes. Vertical error bars are derived from the MCMC parameter uncertainty estimates, and horizontal error bars represent observational uncertainty estimates. Error bars and fit statistics are shown in table 4. .......................................................................................89

Figure 4-1: Map of the analyzed Ameriflux eddy covariance tower locations. Sites with long records (>= 5 years) are selected to assess the effectiveness of the model in reproducing observed interannual variability. Details about the site characteristics are given in Table 1. .................................................................99

Figure 4-2: Schematic diagram of the flow of carbon in the TRIFFID model. Gross primary productivity (GPP), net primary productivity (NPP) and autotrophic respiration ($R_a$) are calculated separately for the needleleaf (NL) and broadleaf (BL) functional types. NPP is allocated to leaf, wood and root pools, which all turn over into a single soil carbon pool. Heterotrophic respiration ($R_{H}$) is a function of soil carbon, soil temperature and soil moisture. NPP is the balance between GPP and $R_a$, and net ecosystem exchange (NEE) is the balance between NPP and $R_{H}$. .....................................................................................103

Figure 4-3: Scan of the log likelihood function over the prior range of the $\theta_{opt}$ (optimal soil water content for heterotrophic respiration) parameter for the WLEF individual optimization. Two maxima are visible, indicating a nonconvexity in this 1D scan. The use of gradient-based optimization algorithms may misconverge to a local solution rather than the desired global solution........................................................................................................................................118
Figure 4-4: Convergence of the SRES algorithm for each individual site optimization. Because SRES is a minimization algorithm, we use the negative of the log likelihood as our objective function. In all five cases, the values of the objective function in runs 2 and 3 are within 0.5% of the value obtained in run 1. Results are similar for the joint optimization (not shown). In all optimizations, the run with the lowest value of the objective function is used to make model predictions.

Figure 4-5: Convergence diagnostics of 4 model parameters the Howland Forest individual site optimization. In all three SRES runs, the soil carbon and Q_{10H} parameters converge to nearly identical values. However, the soil moisture parameters \( \theta_{\text{fac}} \) and \( \theta_{\text{opt}} \), that control the soil moisture dependence of heterotrophic respiration, converge to different values in each of the three runs despite the similar likelihood values illustrated in Fig. 4-4.

Figure 4-6: Mean seasonal cycle of NEE over the available tower record for the observed NEE, model NEE using published parameters, and model NEE using the optimized parameters from the separate optimizations. Each point represents a 20-day average of hourly NEE values. Using published parameters results in a poorly-timed growing season and not enough growing season uptake. The optimized models reproduce the phase and amplitude of the seasonal cycle well, with the exceptions of Howland and WLEF, where too much uptake is modeled in the latter part of the growing season.

Figure 4-7: Mean diurnal cycle of NEE over the growing season (defined as June-August) averaged over all years in the available tower record. Shown are observed NEE (x), model NEE using published parameters (dashed), and model NEE using the optimized parameters from the separate optimizations (solid). Using published parameters results a diurnal cycle that is too small in magnitude. In most cases, estimates of nighttime CO\(_2\) flux (respiration) are near the observed, but estimated daytime fluxes are too small. The optimizations improve this considerably, although still slightly underestimating the amplitude at Harvard, WLEF and UMBS.

Figure 4-8: As in Figure 4-6, except showing a comparison of the separate and joint optimizations to observations. The joint and separate optimization produce similar results in all cases, with a slight degradation in the timing of the seasonal cycle when using the parameters from the joint optimization.

Figure 4-9: Comparisons of observed annual NEE sums to simulated NEE sums computed using published parameters, parameters from the separate optimizations and parameters from the joint optimization. Despite the ability of the model to capture diurnal and seasonal cycles, TRIFFID does not capture interannual variability at any tower in a statistically meaningful way (p < 0.05).
Figure 4-10: Comparison of observed vs. modeled NEE for two cases of the separate assimilation at the WLEF tower: one using the observed soil water content (SWC) and one using the SWC from the simple bucket model in TRIFFID. Using the observed SWC improves the agreement of the model to observations, but most of the interannual variability is still not captured and the range of modeled variability is too small........................................................134

Figure 4-11: Comparison of modeled and observed NEE across sites. Modeled NEE is from the joint assimilation experiment in which one set of parameters is used to model all sites with the exception of soil carbon and leaf nitrogen, which are site-specific parameters. The joint assimilation captures inter-site variability, but fails to capture interannual variability at any site...............................136
LIST OF TABLES

Table 2-1: Estimated model parameters.................................................................25

Table 3-1: Sums of gap-filled NEE (units of gC m$^{-2}$) for each year and for all four seasons of each year. Uncertainty estimates (1σ) are shown for the annual sum and represent the combined estimate of turbulent sampling and gap-filling uncertainties. The correction for level selection bias has already been applied to the annual NEE sum. Separate daytime and nighttime NEE sums for the full year and all four seasons are also shown. ...........................................77

Table 3-2: Correlations between anomalies of NEE sums and (a) soil temperature or (b) soil moisture. Night is defined as all hours when PAR = 0, and day is defined as all hours when PAR > 0. Highlighted values are statistically significant (P < 0.05)..................................................................80

Table 3-3: Parameters used in the interannual NEE model. The best fit solutions are the value of the parameter at which the likelihood function is maximized. The 2.5 and 97.5 percentiles of the posterior parameter PDFs represent the bounds of the 95% posterior confidence intervals derived using Markov Chain Monte Carlo. In all cases, priors were uniformly distributed between the prior min and max. ..........................................................................................83

Table 3-4: Goodness of fit statistics for the interannual model as a function of time of day and time of year. Slopes are determined by performing a linear fit of the model as a function of observed NEE; slopes are less than one in all cases, indicating that the model underestimates interannual variability. Observed means and ranges are determined using gap-filled NEE from the period 1997-2004 excluding 2002. Model bias indicates the deviation of the model mean from the observed mean for the given time of day and year. Highlighted values of $R^2$ are statistically significant (P < 0.05)..........................88

Table 4-1: Characteristics of the five Ameriflux eddy covariance tower sites used in the assimilation of the TRIFFID model. Broadleaf and needleleaf fractions are computed by aggregating species composition data in the literature. Canopy height represents an average over the tower footprint and is set equal for both plant functional types. ............................................................................100

Table 4-2: TRIFFID model parameters optimized in this study. .........................104

Table 4-3: Results of the parameter optimization for each separate optimization at HW (Howland), HV (Harvard), WL (WLEF), MM (Morgan Monroe) and UM (UMBS). Parameters that have converged are shown in bold, and are defined as values that are within 0.5% of each other in all three SRES runs and are not edge-hitting (upper and lower bounds and parameter units are
shown in Table 4-2). For parameters that have not converged, results from the most likely of the three SRES runs are shown.

Table 4-4: Results of the joint parameter optimization, which uses observations from all sites as a data constraint. Parameters that have converged and that are not edge-hitting are shown in bold, (upper and lower bounds and parameter units are shown in Table 4-2).
PREFACE

Chapters 2, 3, and 4 are slight modifications of multi-authored manuscripts that have been submitted or are intended to be submitted for publication. Chapter 2 is currently in press in *Agricultural and Forest Meteorology*, while chapters 3 and 4 have not yet been submitted at the time of publication of this dissertation. Daniel Ricciuto, the author of this dissertation, is the first author on all three of these works. He was primarily responsible for developing hypotheses, data processing, writing and revisions. Coauthors, named in the footnotes on the first page of each chapter, were involved in data collection, funding, methodological guidance and editing.
I would like to thank my adviser, Dr. Kenneth J. Davis, for his patience and insightful guidance as I progressed toward my degree. His active role in my education included not only guiding the progress of my dissertation, but also allowing me to take an active role in interdisciplinary collaboration, proposal writing, paper reviewing, advising of undergraduate students, and in presenting work at a number of conferences. These skills are essential for the development of a successful scientist. I would also like to thank committee member Dr. Klaus Keller, from whom I learned a great deal about statistics, data assimilation, how to read scientific literature effectively, proposal writing, and making presentations. I am also grateful to committee members Ray Najjar and John Wyngaard for providing constructive comments and ideas. All four committee members contributed considerably to my academic education in the classroom and were excellent teachers. Without their encouragement and guidance, this dissertation would not have been possible.

I would also like to thank my colleagues and friends in both the Davis and Keller research groups. I am deeply indebted to Martha Butler, who painstakingly worked with me on seemingly endless revisions of the WLEF flux tower code and provided helpful comments about my research. Other past and present colleagues also provided useful insights and computing help, including Ankur Desai, Ken Craig, Michael Hurwitz, Weiguo Wang, Chuixiang Yi, Kelly Cherrey, Scott Richardson, Tasha Miles, Katie Brennan and Dave McInerney. In addition, the operation of the WLEF flux tower is a
multidisciplinary effort that includes several universities and organizations, and I would like to thank a number of those who participated in this effort: Roger Strand (chief engineer for WLEF-TV), Arlyn Andrews, Peter Bakwin, and Conglong Zhao at NOAA, Ron Teclaw and Dan Baumann at the United States Forest Service, and especially Bruce Cook at the University of Minnesota.

This research was funded in part by the Office of Science (BER), U.S. Department of Energy Terrestrial Carbon Processes program, grant number DEFG02-03ER63681, through the Midwestern Regional Center of the National Institute for Global Environmental Change under cooperative agreement No. DEFG03-90ER61010, and by the U.S. Department of Commerce National Oceanic and Atmospheric Administration (NOAA) Office of Global Programs, grant number NA040AR4310124. Conference, workshop and research travel support was made possible by the NSF Research Collaboration Network (RCN) Grant No. DEB-0130380. Any opinions, findings, and conclusions or recommendations herein are those of the authors and do not necessarily reflect the views of the cooperating agencies.

Finally, I would most of all like to thank my family for their support during my long academic career. My parents, Michael and JoAnn Ricciuto, encouraged my passion for science from an early age. My wife, Kristin Lala Ricciuto, never failed to provide inspiration, loving support and understanding.
Chapter 1

Introduction

1.1 Background

Models of the coupled climate-carbon cycle system vary widely in predictions of future CO₂ sinks (Friedlingstein et al., 2006). The strength of this sink, which consists of uptake through terrestrial ecosystems and the oceans, has important implications for creating emissions targets to reduce the likelihood of dangerous anthropogenic interference in the climate system (IPCC, 2001). Thus, policy decisions must be made in the face of large uncertainties. By assimilating CO₂ flux observations into process-based biogeochemical models, we aim to reduce uncertainty in future predictions.

Atmospheric CO₂ concentrations are currently increasing at an average rate of roughly 1.5 parts per million (ppm) per year, which is about one half of the rate of fossil fuel emissions. The other half of fossil fuel emissions is absorbed by the CO₂ sinks. Unlike fossil fuel emissions, this sink of CO₂ exhibits significant interannual variability, with a 5 GtC range of the 1990s. The magnitude of the total sink is well constrained, as it is determined by the difference between the annual global CO₂ increase and fossil fuel emissions; interannual variations are greater than the range of uncertainty. Annual average atmospheric CO₂ can be measured with high precision (Keeling and Whorf, 2004), and fossil fuel emissions can also be estimated with relatively small uncertainty (Marland et al., 2003). Although inversion models indicate that the terrestrial sink is
smaller than or equal in magnitude to the oceanic sink (Gurney et al., 2002; Rodenbeck et al., 2003b), as much as two thirds of total sink variability has been attributed to the terrestrial biosphere because the strength of the oceanic sink is probably more consistent on an interannual timescale (Bousquet et al., 2000; Rodenbeck et al., 2003a).

The terrestrial carbon sink is driven by two much larger ecosystem fluxes: net primary productivity (NPP) and heterotrophic respiration (R_h). Climate, mainly through temperature, radiation and precipitation, has been shown to be a primary driver of interannual variability in both NPP (Nemani et al., 2002) and respiration (Raich et al., 2002). Regional terrestrial CO_2 fluxes are also significantly correlated with climate indices such as the Southern Oscillation and Arctic Oscillation (Potter et al., 2003). Climate change is thus likely to have a considerable impact on the magnitude of the terrestrial carbon sink in the future. This impact is currently unclear because a number of competing carbon-climate feedback mechanisms cause large uncertainty in predictions of future CO_2 fluxes. For example, two recent studies used different coupled GCM-carbon cycle models to make predictions of CO_2 concentrations and climate for the next 100 years. One model (Cox et al., 2000) predicted a strong positive feedback between respiration and temperature, causing the terrestrial biosphere to switch from a sink to a source of CO_2 after 2050. However, another model maintained the terrestrial biosphere as a sink of carbon through the year 2100 (Dufresne et al., 2002). These models also differ significantly in predictions of interannual variability over the past 50 years. Such studies are computationally expensive, and therefore varying model parameters in order to produce statistical uncertainty estimates is not feasible.
Other researchers have attempted to constrain models of terrestrial carbon cycle variability with global observations of CO$_2$ and climate. Knorr and Heimann (1995) developed the Simple Diagnostic Biosphere Model (SDBM), in which both NPP and $R_H$ are modeled from input variables of Normalized Difference Vegetation Index (NDVI), NPP, temperature, precipitation and solar radiation. Only two adjustable parameters were used: $\varepsilon$ (light use efficiency) and $Q_{10}$ (factor of respiration increase for a 10°C rise in temperature). This model was found to simulate the annual cycle in atmospheric CO$_2$ with a skill several times greater than more complex ecosystem models with many parameters (Heimann et al., 1998). However, this model was not used to predict interannual variability and is of limited use in predicting future fluxes because it required satellite-derived NDVI as input.

Another modeling study involved fitting sixteen ecosystem parameters to produce a reasonably accurate prediction of variations in the strength of the terrestrial carbon sink (Vukicevic et al., 2001). This study demonstrated that lagged temperature effects and nutrient cycling are necessary to include in a model of interannual variability, but did not consider other climate variables or allow for oceanic flux variability. A recent study by Rayner et al. (2005) involved an assimilation of 57 parameters in the Biosphere Energy Transfer Hydrology Model (BETHY) in conjunction with the TM2 transport model using CO$_2$ concentration observations from GLOBALVIEW-CO$_2$ as constraints. The optimized model reproduces the observed global CO$_2$ growth rate reasonably well, but the method used in this study is prone to misconvergence.

We note that none of the data assimilation studies described above incorporated eddy covariance observations. Eddy covariance towers measure continuous fluxes of
CO₂ at half-hourly or hourly timesteps, and they also collect a number of meteorological variables such as air and soil temperature, precipitation and radiation. Currently, there are over 200 such towers spread throughout the world (Baldocchi et al., 2001). Eddy covariance measurements have been used to optimize terrestrial models, improving predictions of carbon fluxes at Harvard Forest (Braswell et al., 2005), Niwot Ridge (Sacks et al., 2006), a site in Kansas (Knorr and Kattge, 2005), and WLEF in northern Wisconsin (Ricciuto et al., 2006). Although model predictions improve from diurnal to seasonal timescales, interannual variability is usually poorly modeled even after parameter optimization.

We believe that eddy covariance observations have been underutilized in data assimilation studies. No multi-site assimilation using eddy covariance data has been attempted to our knowledge. By assimilating flux observations instead of CO₂ concentrations, we eliminate the need for modeled atmospheric transport; this speeds up model evaluations time significantly and remove transport uncertainty. This type of study has the potential to yield important insights about parameterizations in biogeochemical models. Such models usually group together many species and subcategories of vegetation together into a plant functional type (PFT), for example temperate deciduous trees. Each PFT has a specific set of parameters associated with it that is used to make predictions about CO₂ fluxes. Assimilating observations from eddy covariance sites within the same PFT will reveal how coherent these parameters are. Such an analysis will also reveal whether key processes are missing or misrepresented by examining the models’ ability to reproduce observed CO₂ fluxes across a range of spatial and temporal timescales. In particular, we focus on the question of interannual
variability. A better understanding of the effects of climate on interannual variability in terrestrial fluxes and how this is linked with other factors will lead to better forecasts of terrestrial sink strength.

1.2 Outline

Chapter 2 presents a data assimilation study of a simple carbon cycle model coupled to an energy balance model. Markov Chain Monte Carlo (MCMC) is used to calibrate carbon cycle parameters that govern the strength of the terrestrial and ocean carbon sinks. We calculate PDFs of these parameters by assimilating historical observations. Our model maps global anthropogenic carbon emissions to global CO₂ concentrations and fluxes. It can be forced with historical global-average temperature data or can be coupled with a simple energy balance model to predict temperature as a function of CO₂ concentration, in which case there is little interannual variation in the temperature record. This method is used to explore the effects of neglecting autocorrelation in observations of CO₂ and to illustrate the utility of estimating parameter PDFs with observed interannually varying temperature. Additionally, we use the parameter PDFs to make probabilistic predictions of the strength of the carbon sink in the future given emissions scenarios.

In chapter 3, a simple model is used to assimilate observations from a single eddy covariance site: the 447 m WLEF tower in northern Wisconsin. The objective of this study is to analyze seasonal and interannual variability in NEE at this site. We first show that the range of interannual variability is significant compared to random errors in the
observations. Probability density functions (PDFs) of model parameters are then estimated using the MCMC framework and our simple model. Modeled seasonal and annual NEE sums are compared to sums derived from observations. Random and systematic errors in the fluxes are evaluated and compared to model uncertainties.

Chapter 4 is a multi-site data assimilation of eddy covariance data using the TRIFFID model. Five eddy covariance sites in eastern U.S. forests with a combined 37 site-years of data are used. We optimize model parameters individually and jointly for each site in order to evaluate the ability of the model to reproduce observed fluxes from hourly to interannual timescales. Optimized parameters are compared across sites to assess which parameters are the most spatially coherent. We then perform a joint optimization of all five sites, obtaining a single set of model parameters to predict carbon fluxes across all site-years. Variation among sites is assumed to be caused solely by variations in climate forcing, canopy height, PFT distribution, soil carbon and leaf nitrogen content. We then evaluate the ability of the jointly optimized model to reproduce both interannual and cross-site variability. Finally, chapter 5 consists of a summary and conclusions, as well as suggestions for future work.
Chapter 2

A Bayesian synthesis inversion of carbon cycle observations: How can observations reduce uncertainties about future sinks?

2.1 Introduction

About half of anthropogenic CO₂ emissions are currently absorbed by the oceans and the terrestrial biosphere. How these carbon sinks will change in the future is at this time an open question of considerable policy relevance. One common approach to address this question is to employ atmosphere and ocean general circulation models (AOGCMs) that are fully coupled with prognostic terrestrial carbon models (TCMs). Although these state-of-the-art models were developed using the best available mechanistic understanding of climate and the carbon cycle, subtle differences in model parameters and structures translate to large ranges in the predicted strengths of the terrestrial and ocean carbon sinks in the 21st century. For example, the Hadley model used in Cox et al. (2000) predicts a strong positive feedback between climate and the carbon cycle with the terrestrial carbon sink switching to a source in about 2050. In contrast, the IPSL model used in Dufresne et al. (2002) predicts that the terrestrial carbon pool will continue to be an atmospheric carbon sink. An analysis of feedbacks demonstrates that this discrepancy between the models is largely caused by differences in the representation of the terrestrial carbon cycle; the temperature sensitivity of terrestrial

---

1 This chapter is a slightly modified version of a manuscript to be submitted to Global Biogeochemical Cycles: Ricciuto, D.M., Davis, K.J., Keller, K., A Bayesian synthesis inversion of carbon cycle observations: How can observations reduce uncertainties about future sinks?
flux of the Hadley model exceeds the sensitivity of the IPSL model (Friedlingstein et al., 2003). A recent intercomparison of ten fully coupled models in the Coupled Carbon Cycle Climate Model Intercomparison Project (C4MIP) reveals a wide range of carbon-climate feedback strengths, causing sizeable uncertainty about the strengths of both the ocean and terrestrial carbon sinks by the year 2100 (Friedlingstein, in press).

A proper assessment of uncertainty about the carbon sink is necessary for the design of sound, economically efficient carbon management and observation strategies. Although the range of predicted carbon fluxes in the C4MIP intercomparison is a rough measure of model structural uncertainty, these models have not been constrained by the large body of existing observations of CO₂ concentrations and fluxes in a formal statistical sense. Model parameters are often derived from lab-scale studies and may not be applicable to processes on the scale of the model grids (Oreskes et al., 1994). Estimates of predictive uncertainty in individual models are generally restricted to sensitivity analyses of few parameters due to the large computation burdens of these complex models.

One promising approach to constrain parameters and to produce statistically sound estimates of uncertainty about the carbon sink is to assimilate observational data into carbon cycle models. Most data assimilation techniques require a very large number of model evaluations (often exceeding 10²-10⁴), making this computationally infeasible for most fully coupled GCM/carbon cycle models. One groundbreaking study in this regard is Kheshgi and Jain (1999), where a Bayesian technique is applied to estimate 26 global-scale parameters. In this study, observational constraints include globally averaged temperature, atmospheric CO₂ accumulation and carbon isotope data. This
study yielded important insights about the behavior of carbon cycle parameters but is subject to several methodological caveats. For example, they assume normally distributed PDFs and do not consider the potentially valuable annual timeseries of CO$_2$ as an observational constraint because of potential problems associated with autocorrelation. Vukicevic et al. (2001) used global observations of CO$_2$ concentrations to optimize 16 global terrestrial ecosystem parameters with the variational parameter estimation technique, but did not estimate parametric uncertainty. This pioneering study demonstrated the usefulness of adjoint methods but also showed that these methods can fail in the case of nonconvex (global) optimization problems.

Other studies have taken advantage of the NOAA CMDL flask network of CO$_2$ measurements to constrain global or biome-specific ecosystem model parameters. Kaminski et al. (2002) used data from individual CO$_2$ flask stations in order to constrain a small number of parameters in the simple diagnostic Biosphere Model (SDBM) with a Bayesian inversion technique. The fast adjoint version TM2 transport model enabled the mapping of fluxes to flask concentrations in a computationally feasible timeframe. More recently, the Carbon Cycle Data Assimilation System (CCDAS) began as an expansion of this study by using CO$_2$ concentration data and satellite radiation data to optimize 57 biome-specific or global parameters in the Biosphere Energy Transfer Hydrology Scheme (BETHY) terrestrial ecosystem model, also using the adjoint version of the TM2 transport model (Rayner et al., 2005). Both of these studies estimate parametric uncertainty.

Although these studies are a large step forward in constraining process-based models, they are silent on the effects of autocorrelation in residual (observed-modeled)
timeseries and often assume that parameters are normally distributed. The Bayesian studies assume that both the prior and posterior parameters are normally distributed. Neglecting autocorrelation can lead to biased and overconfident parameter estimates (Zellner and Tiao, 1964). Furthermore, the highly nonlinear nature of carbon cycle models is likely to result in non-Gaussian parameter distributions, as evidenced in the cost function of a respiration-temperature sensitivity (Q_{10}) parameter in Kaminski et al. (2002). Parameter distributions in some of these models also display nonconvexity, or multiple optima in the cost function (Rayner et al., 2005; Vukicevic et al., 2001). Assuming that parameters are normally distributed often forces the consideration of physically impossible solutions (e.g. negative diffusivities) because of the symmetry of the tails of the Gaussian distribution. Assuming normality and neglecting autocorrelation are both likely to cause biased estimates of parametric uncertainty. In nonconvex problems, using gradient-based optimization methods can cause misconvergence to local rather than global optima.

The Markov Chain Monte Carlo approach (MCMC) is an alternative method to assimilate CO₂ concentration and flux data. A major advantage of MCMC is that it is able to recover joint parameter probability density functions (PDFs) without parametric PDF assumptions (e.g., normally distributed). The effects of autocorrelation can be examined in a straightforward manner using this technique. MCMC is computationally expensive compared to the methods used above and generally requires greater than 10⁴ model evaluations, limiting the complexity of models that can be used with this technique and making most models that require atmospheric transport infeasible. This technique was introduced into the field of ecological modeling by Harmon and Challenor (1997),
and has been used to optimize terrestrial ecosystem models at individual eddy covariance sites (Braswell et al., 2005; Knorr and Kattge, 2005; Ricciuto et al., 2005), and to assimilate paleo-data in a climate model (Hargreaves and Annan, 2002).

Here we apply MCMC to a global carbon cycle model to calibrate carbon cycle parameters that govern the strength of the terrestrial and ocean carbon sinks. We use a simple but mechanistically sound and computationally efficient carbon cycle model to calculate PDFs of three carbon cycle parameters by assimilating historical observations. This zero-dimensional model maps global anthropogenic carbon emissions to global CO2 concentrations and fluxes. It can be forced with historical global-average temperature data or can be coupled with a simple energy balance model to predict temperature as a function of CO2 concentration, in which case there is little interannual variation in the temperature record. This method is used to explore the effects of neglecting autocorrelation in observations of CO2 and to illustrate the utility of estimating parameter PDFs with observed interannually varying temperature. Additionally, we use the parameter PDFs to make probabilistic predictions of the strength of the carbon sink in the future given emissions scenarios.

In the future, uncertainty about the strength of the carbon sink may be reduced by additional observations. The power of these observations to reduce uncertainty depends upon the observation type, timing and error. We estimate the utility of additional carbon cycle observations to reduce future carbon sink uncertainty by assimilating hypothetical observations of global terrestrial and oceanic carbon fluxes in the model.
2.2 Methods and Data

2.2.1 Carbon cycle model

The model consists of a simple climate module and a carbon cycle module, which is subdivided into terrestrial and ocean carbon cycle models (Fig. 2-1). It is a zero-dimensional model that estimates global CO₂ concentrations and fluxes as functions of anthropogenic emissions. The climate module and ocean carbon cycle model are taken from the nonlinear impulse-response model of the coupled carbon cycle-climate system (NICCS), as described by Hooss et al. (2001). The climate module is a simple pulse-response energy balance model calibrated against the Hamburg AOGCM, using only CO₂ as a forcing variable with a logarithmic dependence on concentration to estimate globally averaged temperature. The climate sensitivity (hypothetical equilibrium temperature change for a doubling of CO₂) is set to 3.4°C, which is the expected value estimated by Andronova and Schlesinger (2001). The model can be run in coupled mode (using the climate module), or the temperatures can be specified externally using the historical record. Future predictions, of course, must be made using the coupled mode.
The ocean model is the 4-box diffusion model described in Hooss et al. [2001]. The first box consists of both the atmosphere and the surface layer of the ocean, which are equilibrated at every timestep. Layer depths and diffusion coefficients were calibrated against the inorganic version of the Hamburg Ocean Carbon Cycle model (HAMOCC). Equilibration between the surface layer and the atmosphere requires balancing oceanic $p_{CO_2}$ with globally averaged $CO_2$ concentrations. For relatively small changes in dissolved inorganic carbon (DIC), the change in $p_{CO_2}$ can be estimated using the following relationship (Revelle and Suess, 1957):

$$Q_{10} \eta \theta \beta$$

Figure 2-1: Model schematic diagram showing the climate and carbon cycle modules, where the carbon cycle module is subdivided into terrestrial and ocean components. Constraints are illustrated by shaded ovals, while emissions forcing are white ovals. Observations are used to constrain four model parameters.
We use a globally averaged value of the buffer (Revelle) factor $\xi$, which is calculated as a function of DIC and temperature using the method described by Yi et al. (2001). Because $\xi$ depends upon temperature and DIC in a nonlinear way, we must use an iterative method to calculate $p_{\text{CO}_2}$ from DIC. For speed of calculation, the fractional partitioning of anthropogenic carbon between atmospheric and surface layer ocean carbon is implemented as a lookup table. The model neglects the effects of changes in ocean circulation and biota on carbon uptake. We estimate the transfer velocity across the thermocline $\eta$ using the MCMC data assimilation process, but do not vary the other ocean model parameters. This is in accordance with Sarmiento et al. (1992), who show that thermocline diffusivity is key in controlling the magnitude of oceanic uptake of anthropogenic carbon.

The terrestrial model was adapted from Siegenthaler and Oeschger (1987), which forms the basis of the BERN carbon cycle model (Joos et al., 1996). It consists of 4 global carbon pools: leafy vegetation, living wood, detritus, and soil carbon. Preindustrial pool sizes are shown (Fig. 2-2). Net primary productivity (NPP) is partitioned at every timestep into the leaf and wood carbon pools. Preindustrial NPP ($\text{NPP}_0$) is 60 GtC yr$^{-1}$, and this is modified by carbon fertilization using a logarithmic relationship (Joos et al., 1996):

$$NPP(\text{CO}_2, t) = \text{NPP}_0 (1 + \beta \frac{\ln([\text{CO}_2](t))}{\ln([\text{CO}_2]_0)})$$

\[ \text{(2.2)} \]
where $\beta$ is the carbon fertilization parameter (dimensionless), $[\text{CO}_2](t)$ is the globally averaged carbon dioxide concentration at time $t$ in parts per million (ppm), and $[\text{CO}_2]_0$ is the preindustrial carbon dioxide concentration (280 ppm). For the purpose of this study, we neglect temperature effects on NPP.

Given preindustrial conditions, the modeled terrestrial carbon cycle is in equilibrium; heterotrophic respiration balances $\text{NPP}_0$, and the net flux from terrestrial ecosystems to the atmosphere is zero. A portion of the leaf and wood pools are converted

---

**Figure 2-2:** Schematic of the terrestrial carbon cycle model. NPP is a function of atmospheric CO$_2$ concentration, and is allocated to leaf and wood pools. These pools turn over at specified rates into the detritus pool and soil carbon pools. These pools then turn over in the atmosphere at a rate dependent on the global mean temperature. Preindustrial pool sizes are given in bold.
to detritus and soil carbon at each timestep, and a portion of detritus is converted to soil carbon. Heterotrophic respiration of CO$_2$ to the atmosphere occurs from the detritus and soil carbon pools, represented by a loss of a fraction of each pool every timestep. Heterotrophic respiration rates ($R_H$) are modified from preindustrial values using the $Q_{10}$ relationship:

$$R_H(\Delta T, t) = \left[ S_1(t) + S_2(t) \right] ^{\left(\frac{\Delta T-10^\circ C}{10^\circ C}\right) Q_{10}} \tag{2.3}$$

where $S_1$ is the size of the detritus pool in GtC, $S_2$ is the size of the soil carbon pool in GtC, $\Delta T$ is the global temperature deviation from the preindustrial mean ($^\circ C$), and $Q_{10}$ is a dimensionless parameter controlling the sensitivity of respiration to temperature. A $Q_{10}$ of 2, for example, implies a doubling of global respiration rates for a $10^\circ C$ increase in temperature given a constant carbon pool size. Note that respiration rates depend not only on $Q_{10}$ and temperature at time $t$, but also on the sizes of the respiring carbon pools. This results in an increase in respiration through the indirect effects of CO$_2$ fertilization, which causes equilibrium pool sizes to grow larger as a function of increasing CO$_2$ concentrations. For simplicity, we assume that the same $Q_{10}$ values apply to the detrital and soil carbon pools. We also assume that changing values of CO$_2$ and temperature affect neither the partitioning of NPP between the leaf and wood pools, nor the partitioning of carbon within pools.
2.2.2 Model forcings

The calibration period of the model is 1850-2004. Future predictions are made using specified emissions or stabilization scenarios from 2005-2100. In 1850, terrestrial NPP and pool sizes are set to their preindustrial values, the global temperature anomaly $\Delta T$ is set to zero, and the atmospheric concentration is set to 280 ppm. The model is then in equilibrium with no net ocean-atmosphere or net terrestrial-atmosphere carbon fluxes. The model is forced with anthropogenic emissions, including emissions from fossil fuel burning, cement production and land-use change. Global estimates of carbon emissions from fossil fuel burning and cement production from 1850-2002 are from Marland et al. (2005). We extend this estimate to 2003 and 2004 by linear extrapolation from the previous 10-year period. Land-use emissions from the period 1850-2000 are estimates from Houghton et al. (2003), with linear extrapolation to 2001-2004 from the previous 10-year period. Additionally, the model may be forced with a historical temperature record when it is not coupled with the energy balance model. In these cases, we use the globally averaged temperature dataset covering the period from 1856-2004 from Jones et al. (2005). We subtract the mean of the period over the first 30 years of the record, so that the average temperature over this period is assumed to equal the preindustrial mean (where $\Delta T = 0$). We also set $\Delta T$ to zero during the period 1850-1855, when global temperature data were not available. We neglect uncertainties in the temperature and emissions data.
2.2.3 Observational Constraints

The model is constrained by observations of CO$_2$ concentrations and estimates of global ocean fluxes. CO$_2$ concentration constraints include data from the Law Dome ice core between 1850-1959 (Etheridge et al., 1996) and from the Mauna Loa Observatory from 1960-2004 (Keeling and Whorf, 2005). Both sites were chosen as proxies for the global average. The Law Dome dataset contains 17 points representing different samples of air in the ice core with different ages. We assume that the observation error of each estimate is normally distributed with a standard deviation of 8 ppm. The mean air age of the sample was taken as the year of observation for the purpose of model evaluation. The annual Mauna Loa CO$_2$ contains 45 data points, and the observation error of each measurement is assumed to be 0.4 ppm. The error estimates for the Law Dome and Mauna Loa data are higher than the published values of instrumental error because they include estimates of process error. The reasoning behind these choices is discussed below.

A number of independent estimates of oceanic sink strength exist that can be used to constrain our model. The World Ocean Circulation Experiment (WOCE) in the 1990s included an extensive survey of dissolved inorganic carbon (DIC) and related tracers. Based on these observations, Sabine et al. (2004), estimate a cumulative oceanic sink for anthropogenic CO$_2$ of 118 +/- 19 GtC from 1800 to 1994. We assume that before 1850, the oceans were in preindustrial equilibrium. We hence take Sabine’s estimate as a constraint on the cumulative ocean sink from 1850-1994. This may be a slight overestimation of uptake because of possible anthropogenic emissions that were absorbed
before 1850 (c.f. Ruddiman, 2003). A second constraint on the strength of the oceanic carbon sink arises from the chlorofluorocarbon (CFC) dataset, which was combined with measured DIC to estimate average annual uptake over the decades of the 1980s and the 1990s. McNeil et al. (2003) estimate an average annual oceanwide uptake of 1.6 +/- 0.4 GtC yr\(^{-1}\) over the 1980s, and 2.0 +/- 0.4 GtC yr\(^{-1}\) over the 1990s. To compare the model output to these estimates, modeled annual ocean fluxes were averaged over each decade.

### 2.2.4 Data assimilation technique

Markov Chain Monte Carlo (MCMC) is a powerful method to assimilate observations into nonlinear models that are relatively fast and have a modest number of parameters. MCMC simulates direct draws from a joint probability density function, making no assumptions about the shape of this distribution. Prior information about parameters can be included using a Bayesian framework; using diffuse uniform priors recovers the maximum likelihood approach in a frequentist sense. We use the Metropolis-Hastings algorithm (Hastings, 1970; Metropolis et al., 1953), which has shown to be relatively robust (Qian et al., 2003).

In the Metropolis-Hastings algorithm, the posterior probability \(p_{\text{post}}\) of a sample from the joint parameter distribution \(\theta_k\) is a function of the prior probability \(p_{\text{prior}}\) and the observations \(x\) according to:

\[
p_{\text{post}}(\theta_k | x) = \frac{L(x | \theta_k)p_{\text{prior}}(\theta_k)}{\sum_{i=1}^{N} L(x | \theta_i)p_{\text{prior}}(\theta_i)}
\]
The denominator on the right-hand side of this equation is a normalization factor such that the sum of posterior probabilities over all possible sets of parameters is equal to one. The construction of the Markov chain only requires knowledge of probability and likelihood ratios, so that we may ignore this normalization factor during this step. The function $L$ is the likelihood of observing all of the observations $x$ given the set of parameters $\theta_k$. The likelihood for observations with independently and identically distributed (IID) error drawn from a normal distribution is given by:

$$L(x | \theta_k) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{1}{2} \left[ \frac{f(\theta_k,t_i) - x_i}{\sigma_i} \right]^2 \right)$$

The likelihood of a set of $n$ observations in $x$ is the product over the likelihoods of individual observations $x_1, x_2, \ldots, x_n$ at times $t_1, t_2, \ldots, t_n$. We assume that each observation is sampled independently from a normal distribution with a mean $x_i$ and standard deviation $\sigma_i$ representing the variability due to observation noise and internal variability that is not captured by the model. Observation error may differ in magnitude from one observation to the next. This likelihood is a product taken over multiple observations and observation types such as CO$_2$ concentrations and ocean fluxes.

This formulation of the likelihood function follows previous studies using MCMC that assume IID and normally distributed variability (Braswell et al., 2005; Hargreaves and Annan, 2002). This can be a useful and reasonable approximation if the observation errors are large and if the process noise is indeed IID. However, many environmental timeseries show statistically significant autocorrelation (e.g. Lund et al., 1995). In this case, a more refined likelihood function has to be used. This can be illustrated
specifically for the annual CO2 concentrations as observed at Mauna Loa from 1960-2004, which displays significant autocorrelation that cannot be reproduced by our simple model. A different form of the likelihood function is used to test the hypothesis that neglecting autocorrelation in the residuals (modeled – observed) of annual CO2 concentration causes biased and overconfident parameter estimates. This form does not assume that the observations are sampled independently, but that the error term displays significant lag-1 autocorrelation. This total error term represents both process error and observational error. We assume that the model is a perfect representation of the true globally averaged CO2 concentration $y$:

$$ y_t = f(\text{emissions}, \theta, t) \quad 2.6 $$

The variable $x_t$ representing observed CO2 concentrations includes an autocorrelated error term $u_t$ according to:

$$ x_t = f(\text{emissions}, \theta, t) + u_t $$

$$ u_t = \rho u_{t-1} + \epsilon_t \quad 2.7 $$

The $\epsilon_t$ term represents IID error, which is assumed to be independent and normally distributed with a mean of zero and a standard deviation $\sigma$. Given these assumptions, the likelihood function is given by Zellner and Tiao (1964):

$$ L(x | \theta_k) = \exp \left\{ -\frac{1}{2\sigma^2} (f(\theta_k, t_0) - x_0 - M)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (f(\theta_k, t_i) - x_i - \rho f(\theta_k, t_{i-1}) - \rho x_{i-1}) \right\} \quad 2.8 $$
This represents the likelihood of observing $x$, which is a vector of autocorrelated observations with $n$ timesteps, given a set of parameters $\theta$. $\rho$ is the lag-1 autocorrelation parameter (correlation obtained when the timeseries is lagged by one year). We estimate the PDF of this parameter in the same manner as the carbon cycle parameters. These assumptions also require that the uncertainty in the mean of the observation for the first timestep ($t=0$) be specified as an unknown parameter $M$. Unlike Eq. 2.5, this form of the likelihood function assumes that all observations have the same standard deviation $\sigma$.

To construct the joint parameter PDF, we must sample many sets of parameters $\theta$. In a Bayesian framework, prior information is combined with the likelihood to produce a posterior probability using Eq. 2.4. One simple approach to this task would be to compute likelihoods at random points over the range of the prior distribution and combine this information to produce a PDF. However, most of the posterior probability mass may be concentrated in a small area of the prior range, making this sampling method computationally inefficient. MCMC makes use of information about the gradient of the likelihood function in order to preferentially sample in regions where the posterior probability is high. Because MCMC uses gradient-based information, it is best suited for convex problems with a single maximum of the likelihood function.

To construct our Markov chain, we begin with an initial parameter guess $\theta_i$. This initial guess is unimportant in determining the posterior parameter PDFs, but does affect the length of burn-in time required for the Markov chains to become stationary (see below). We calculate the likelihood $L_i$ of the observations $x$ given the initial guess $\theta_i$, or $L_i(x|\theta_i)$. Next, we advance from the initial guess $\theta_i$ to a candidate point $\theta_c$ using a
uniform proposal distribution, meaning that it is equally likely to step in any direction in the n-dimensional parameter space where n is the number of model parameters. The maximum step size for each parameter is an algorithm-specific parameter that is specified \textit{a priori}. We then compute the likelihood of the candidate point $L_c(x|\theta_c)$. The algorithm that determines whether or not to accept the candidate point as the next step in the Markov chain:

$$\frac{L(x|\theta_c) \cdot p_{\text{prior}}(\theta_c)}{L(x|\theta_i) \cdot p_{\text{prior}}(\theta_i)} > U$$

In this method, the posterior probability ratio of the proposal step to the initial value must be greater than U, which is a uniform random variable that ranges between 0 and 1. Therefore, if the posterior probability of $\theta_p$ is greater than that of $\theta_i$, the proposal step will always be accepted. If the posterior probability of $\theta_p$ is less than that of $\theta_i$, the proposal step will be accepted a probability equal to the ratio of posterior probabilities $p(\theta_p)/p(\theta_i)$. Setting U = 1 would assure that the Markov chain always moves up-gradient towards the maximum likelihood solution (or a local maximum if the likelihood function contains multiple maxima). When U is allowed to vary between 0 and 1, this allows for the possibility for the Markov chain to move down-gradient in order to sample the posterior distribution entirely. If the step is accepted, the candidate point becomes the new initial point and the process is repeated until the Markov chain has converged.

After a sufficiently long iteration (referred to as the “burn-in” period), the chain becomes stationary. One useful property of the MCMC algorithm is that this stationary distribution converges to the joint parameter posterior PDF. The length of the burn-in period can be minimized by choosing $\theta_i$ that is close to the maximum likelihood solution.
However, the length of the burn-in period and the length of the entire Markov Chain required to simulate the PDF are problem-dependent and difficult to determine. A choice of $10^6$ simulations with a burn-in of $10^5$ results in numerically stable results for the considered problem. We choose proposal step sizes that result in acceptance rates between 25 and 50% following previous studies (Harmon and Challenor, 1997). Removing the burn-in values yields a Markov chain of at least $2 \times 10^5$ parameter sets that are used to construct parameter estimates and uncertainties. This value is considerably lower than the total function evaluations because many proposal steps are not accepted. A comparison among multiple chains with different starting points indicates that a length of $2 \times 10^5$ is sufficient to converge to the joint posterior PDFs (results not shown).

2.2.5 Hypotheses and experimental setup

Carbon cycle parameter values are chosen from the literature, and the value of the lag-1 autocorrelation $\rho$ is set to 0 in accordance with the null hypothesis that there is no autocorrelation in the residuals of annual CO$_2$ concentration data (Table 2-1). Parameter prior distributions are nearly uninformative (uniformly distributed over a large range) where possible, but nonphysical values of parameters are excluded. The thermocline transfer velocity $\eta$ prior ranges from 0 to 1000 m yr$^{-1}$ to exclude nonsensical negative values. The carbon fertilization parameter $\beta$ ranges from 0 to 1000 so that increased atmospheric CO$_2$ concentrations must increase NPP. Similarly, $Q_{10}$ ranges from 1 to 1000 so that increased temperature must have the effect of increasing respiration. The upper limit of 1000 for all three of these parameters is an arbitrary value chosen to be
large enough that the Markov chain is never actually constrained by this bound. The lag 1 autocorrelation parameter $\rho$ is a correlation coefficient and is allowed to range between -1 and 1.

Three separate model calibrations are performed: (a) base case: with observed interannually varying temperature and accounting for autocorrelation in the likelihood function as in equation (4), (b) neglecting autocorrelation: with observed temperature and neglecting autocorrelation in the likelihood function as in equation (6), and (c) smooth temperature: with modeled (smooth) temperature and accounting for autocorrelation in the likelihood function. All assimilations use the same number of model evaluations, initial guesses, step sizes and prior distributions for model parameters except that in the second experiment, the autocorrelation parameter $\rho$ is held constant at zero. Three sets of joint parameter PDFs are then obtained. Comparisons among these calibrations are used to test the hypotheses that (a) Neglecting autocorrelation results in biased and overconfident parameter estimates, and (b) Forcing the model with observed temperature is a valuable constraint on the terrestrial carbon cycle parameters.

Table 2-1: Estimated model parameters

<table>
<thead>
<tr>
<th>Parameter (units)</th>
<th>Starting value</th>
<th>Prior range</th>
<th>Max likelihood</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{10}$ (dimless)</td>
<td>2.0</td>
<td>[1, 1000]</td>
<td>1.5</td>
<td>[1.1, 2.4]</td>
</tr>
<tr>
<td>$\beta$ (dimless)</td>
<td>0.29</td>
<td>[0, 1000]</td>
<td>0.62</td>
<td>[0.48, 0.76]</td>
</tr>
<tr>
<td>$\eta$ (m yr$^{-1}$)</td>
<td>17</td>
<td>[0,1000]</td>
<td>27</td>
<td>[17, 61]</td>
</tr>
<tr>
<td>$\rho$ (dimless)</td>
<td>0</td>
<td>[-1, 1]</td>
<td>0.82</td>
<td>[0.74, 0.99]</td>
</tr>
</tbody>
</table>
The Markov chain generated from the base case simulations is used for predictions and additional analysis. We predict the magnitude of the allowable CO₂ emissions through the year 2100 under the S550 ppm IPCC stabilization scenario from the years 2005-2100. Using a fixed concentration rather than anthropogenic emissions as input requires a modified model construction. For each year, the temperature deviation ΔT is calculated given the specified CO₂ concentration, and the strength of the carbon sink is then estimated given the model parameters, CO₂ concentration, ΔT, and an initial estimate of allowable emissions. The quantity allowable emissions is defined as the amount of anthropogenic emissions (land use + fossil fuel) necessary to stay on the stabilization scenario concentration trajectory, given the strength of the carbon sink. Because the carbon sink strength depends on the amount of anthropogenic emissions at each timestep, we must use an iterative process to equilibrate this sink with allowable emissions. Each member of the Markov Chain is used to produce a prediction of allowable emissions through the year 2100. 95% confidence intervals are then constructed from this probabilistic information.

Finally, we test the hypothesis that given our model structure, a single hypothetical “observation” of annual terrestrial flux reduces the 95% range of sink uncertainty more than an observation of annual ocean flux with the same observational uncertainty. To this end, we assimilate an observation of annual ocean flux in the year 2004 that is 2.4 GtC yr⁻¹, which is equal to the maximum likelihood estimate from the base case scenario. Following McNeil et al. (2003), we assume an observation error of 0.4 GtC yr⁻¹. We then conduct a second experiment in which we assimilate an observation of terrestrial flux in the year 2004 that is 2.6 GtC yr⁻¹, which is again equal to
the maximum likelihood estimate from the base case scenario. We also assign this observation an uncertainty of 0.4 GtC yr\(^{-1}\). To compute parameter PDFs with these hypothetical observations, we rerun the Markov chains using the same methodology with the base-case conditions. The resulting parameter estimates and predictions are then compared with the base-case.

### 2.3 Results and Discussion

#### 2.3.1 Base case parameter distributions

Observations of CO\(_2\) concentrations and fluxes were assimilated into our simple model to construct posterior PDFs of 4 parameters for three experimental cases (Fig. 2-3). The base-case posterior PDFs for the respiration temperature sensitivity \(Q_{10}\), the thermocline diffusivity \(\eta\), and the lag-1 autocorrelation parameter \(\rho\) are not normally distributed (\(p < 0.05\), Lilliefors hypothesis test). Data assimilation methods that make the assumption of normality would have forced the consideration of physically unreasonable values of the respiration temperature sensitivity \(Q_{10}\) and the autocorrelation parameter \(\rho\). The thermocline diffusivity \(\eta\) is reasonably well-described by a log-normal distribution.

The posterior \(Q_{10}\) distribution encompasses most results from global models and lab experiments, although most of the probability mass is concentrated at the low end of previous estimates. Raich and Schlesinger (1992) compiled estimates of \(Q_{10}\) from individual study sites in the literature and reported a median value of 2.4 with a range
Our PDF is sharper and is concentrated in the lower half of this range with a maximum likelihood estimate of 1.5. This data-based estimate is lower than the value adopted in many coupled climate-carbon models (Cox et al., 2000; Zeng et al., 2004).

Using a global $Q_{10}$ relationship to model the response of heterotrophic respiration to...
climate change is a parsimonious approach that neglects both regional effects and other climate variables that may be important. The effects of precipitation have not been considered, although precipitation and soil moisture have been incorporated into other regional-scale models of respiration (Raich et al., 2002; Reichstein et al., 2003). The relationship between global precipitation and respiration is not well known.

The mode of the estimated fertilization parameter ($\beta$) is larger than most estimates in the literature, implying a large response of NPP to increasing CO$_2$ concentrations in this model. This conclusion is contingent on the model structure and the adopted estimate of land-use carbon emissions. Choosing an alternative land-use emissions estimate other than the adopted Houghton et al. (2003) would alter the PDF of $\beta$. Jain et al. (2005), for example, conclude that land-use emissions estimates derived from Houghton (2003) are considerably larger than those based on the data from Ramankutty and Foley (1999). Different ecosystem models produce a significant spread in estimates of land-use emissions even when forced with the same land-use change dataset (McGuire et al., 2001). Because the oceanic sink is constrained, a land-use emission uncertainty would translate into uncertainty in the terrestrial carbon cycle parameters, especially $\beta$.

In the Bern carbon cycle model, which forms the basis for our terrestrial model and uses the same formulation for carbon fertilization, a value of $\beta = 0.287$ was used to balance an estimated deforestation source of 1.1 GtC yr$^{-1}$ in the 1980s. (Joos et al., 1996). Our maximum likelihood estimate of $\beta$ is more than twice this value, but within the range of previously used values (Kicklighter et al., 1999). This result is primarily driven by the fact that Houghton (2003) estimates a larger anthropogenic land-use emission flux of
nearly 2.0 GtC yr\(^{-1}\) in the 1980s, which is compensated by a higher \(\beta\) in the assimilation. In addition, our model requires a larger value of \(\beta\) to balance the respiration increase resulting from increased temperature, which is not considered in the Bern model (Joos et al., 1996). The estimate of \(\beta = 0.6\) implies that NPP in the year 2004 has increased by 11 GtC yr\(^{-1}\) (18\%) since preindustrial times.

The global aggregate modeling approach allows for a rigorous statistical analysis with small computational demands but limits the number of considered processes. Changes in regional temperatures, precipitation patterns and nitrogen fertilization may be important drivers of the terrestrial carbon sink. For example, Bruno and Joos (1997) conclude that carbon fertilization is insufficient to explain the biospheric sink. Caspersen et al. (2000) estimate that growth enhancement due to carbon fertilization is negligible in the upper Midwest region of the United States, and that forest regrowth is the primary reason for the carbon sink in this region. However, this may not be true of all regions, and the forest inventory data used in that analysis may lack the precision necessary to make the conclusion that carbon fertilization is negligible (Joos et al., 2002).

The mode of the estimated thermocline diffusivity (\(\eta\)) is approximately 25.0 m yr\(^{-1}\). This is higher than the 16.9 m yr\(^{-1}\) used in the original version of NICCS that was calibrated against the Hamburg Ocean Carbon Cycle model (Hooss et al., 2001). There is a wide distribution of values, as the 95\% confidence interval ranges from 17.3 m yr\(^{-1}\) to 63.2 m yr\(^{-1}\). This translates to a large uncertainty in the strength of the ocean carbon sink despite the observational constraints. Note that within the adopted model structure, the observational constraints of Sabine et al. (2004) and McNeil et al. (2003) slightly
contradict each other, which magnifies the uncertainty associated with the ocean carbon sink. This is discussed further below.

The assimilation produces significant correlations ($p < 0.05$) in two pairs of model parameters (Fig. 2-4). There is a significant positive correlation between (i) the respiration temperature sensitivity ($Q_{10}$) and the carbon fertilization parameter ($\beta$) and (ii) the carbon fertilization parameter ($\beta$) and the thermocline diffusivity ($\eta$). As discussed above, the atmospheric CO$_2$ budget provides a strong constraint on the total sink strength but is silent on partitioning between the terrestrial and ocean sinks. The considered ocean sink observations constrain the terrestrial sink, because the total sink is given by the atmospheric CO$_2$ budget. Uncertainty in the observational constraints on the ocean sink hence directly translates into uncertainty in the partitioning of the terrestrial and ocean carbon sinks. As a result, the carbon fertilization factor ($\beta$) is positively correlated with the strength of the terrestrial sink, and the thermocline diffusivity ($\eta$) is positively correlated with the strength of the ocean carbon sink. The inverse correlation between the carbon fertilization factor $\beta$ and the thermocline diffusivity $\eta$ results because high values of the carbon fertilization factor and low values of thermocline diffusivity match the observational constraints in a similar way as low values of the carbon fertilization factor and high values of thermocline diffusivity.

Our analysis does not consider observational constraints on terrestrial gross fluxes (NPP and respiration) as these values are typically the results of model analyses. A data synthesis of the global terrestrial carbon uptake (similar to the oceanic constraints) would likely be a strong constraint on carbon flux partitioning. The observed positive
The correlation between respiration temperature sensitivity ($Q_{10}$) and the carbon fertilization factor ($\beta$) is expected because the carbon fertilization factor controls the magnitude of NPP and the respiration temperature sensitivity $Q_{10}$ controls the magnitude of respiration. High values of respiration and high values of NPP produce a similar net terrestrial sink as low values of respiration and low values of NPP. Our analysis suggests no significant

Figure 2-4: Scatter plots representing the joint 95% confidence bounds of (a) respiration sensitivity $Q_{10}$ and carbon fertilization factor $\beta$, (b) $Q_{10}$ and thermocline diffusivity $\eta$, (c) $\beta$ and $\eta$, (d) $\eta$ and the autocorrelation parameter $\theta$, (e) $\beta$ and $\theta$, and (f) $\eta$ and $\theta$. 
correlation between respiration temperature sensitivity ($Q_{10}$) and thermocline diffusivity ($\eta$), or between the lag-1 autocorrelation parameter ($\rho$) and any of the other three parameters.

### 2.3.2 Effects of neglecting autocorrelation

Neglecting autocorrelation in the likelihood function changes the character of parameter PDFs (Fig. 2-3). Neglecting autocorrelation biases the mode of the posterior $Q_{10}$ value by about 30% and results in an artificial tightening of the 95% confidence limit by more than 60%. Similarly, the carbon fertilization factor $\beta$ is shifted toward lower values with a maximum likelihood estimate that is 15% lower and a 95% confidence interval that is 40% tighter when neglecting autocorrelation. The estimate of $\eta$ is less affected. Neglecting autocorrelation violates an important assumption used in constructing the likelihood function and leads to strongly overconfident and biased conclusions.

Autocorrelation in the Mauna Loa timeseries is reasonably well removed by using an AR1 model to whiten the residuals. Before whitening, the lag-1 partial autocorrelation coefficient $\rho$ of the residuals is significant at the 95% level; higher order coefficients are not statistically significant. The maximum likelihood estimate of $\rho$ is 0.83, indicating that the data points are not independent. The residuals are whitened using the approximation:

$$\epsilon_t = (y_t - x_t) - \rho(y_{t-1} - x_{t-1})$$

2.10
where $y_t$ is the modeled CO$_2$ concentration at time $t$, $x_t$ is the observed CO$_2$ concentration at time $t$, and $\varepsilon_t$ is an independent and normally distributed random variable. The partial autocorrelation function shows no evidence of remaining autocorrelation. The standard
deviation of these whitened residuals is 0.43, consistent with the value of 0.4 that we assumed for observational error.

2.3.3 Effects of interannually varying temperature

Using interannually varying observed temperature data to optimize model parameters in the base case results in more sharply constrained parameters, which can be seen by comparing the base-case PDFs with those derived using smooth temperature (Fig. 2-3). Because the temperature sensitivity of respiration is the only mechanism in the model capable of responding at interannual timescales, interannually varying temperature constrains the $Q_{10}$ parameter more than the other parameters. The respiration temperature sensitivity ($Q_{10}$) estimated based on smoothed temperature forcing is much more diffuse with a 95% confidence range 5 times as wide as in the base case. Because $Q_{10}$ and the carbon fertilization parameter $\beta$ are strongly correlated, this constraint on $Q_{10}$ is also effectively a constraint on $\beta$. The 95% confidence range of $\beta$ is more than twice as wide when using smooth temperature; using interannually varying temperature constrains $\beta$ to the lower half of this range. There is almost no difference in the thermocline transfer velocity $\eta$ between the base case and the smooth temperature case because the ocean sink does not respond at interannual timescales in our model, and this sink is already constrained by observations.
2.3.4 Model performance

Modeled CO₂ concentrations and fluxes agree reasonably well with observational constraints (Fig. 2-6). The largest deviations between modeled and observed CO₂ concentrations occur between the period from 1900-1950, during which Law Dome CO₂ measurements are used as constraints. CO₂ concentration growth rates are underestimated during the period 1900-1930, while these growth rates are overestimated during the period from 1930-1950. Etheridge et al. (1996) note that during the 1930-1950 period, there is an apparent decline in the CO₂ growth rate to near zero that cannot be explained by the fossil fuel or land-use emissions alone. They hypothesize that an increased biospheric sink, either through reduced respiration or increased uptake, is responsible for this prolonged growth rate anomaly. This possible enhanced biospheric uptake is not reproduced by our model. Because of the poor model performance during this period, a value of 8 ppm is assumed for the error in the Law Dome ice core constraint, which is several times larger than the uncertainty estimated by Etheridge et al. (1996).

The model fits the long-term trend of the Mauna Loa data remarkably well (Fig. 2-6a). Residuals between the base-case model and observation are seen more easily in Figure 3. The unwhitened residual is mostly negative before 1991 and mostly positive thereafter. The low growth rate anomalies between the years of 1991-1992 and 1992-1993 are not captured by the model. These low CO₂ growth rates, which followed the Pinatubo eruption in 1991, occurred despite a reduction in high-latitude NPP, and are
therefore likely to be the result of a reduction of heterotrophic respiration in boreal regions that is even larger than the reduction in NPP (Lucht et al., 2002). Although there was a low global temperature anomaly associated with the Pinatubo eruption, our model was not successful in reproducing the CO₂ growth rate anomaly through the global Q₁₀.
mechanism. Most of the other interannual variability between the years 1960 and 2004 is also not captured by our model. One reason for this may be that our simple model has no explicit representation of the El-Nino Southern Oscillation (ENSO). ENSO is hypothesized to be one of the primary drivers of interannual carbon sink variability; model analyses that can resolve spatial variability in fluxes suggest that adding this driver would likely improve our model (Zeng et al., 2005).

Although there is little uncertainty associated with modeled CO₂ concentrations, there is considerably more uncertainty associated with the modeled terrestrial and ocean carbon sinks (Fig. 2-6). In 2004, the 95% confidence range of the terrestrial sink was between about 1.8 and 3.2 GtC yr⁻¹; the 95% confidence range of the oceanic sink was between about 1.7 and 3.1 GtC yr⁻¹. As a result of simplifying assumptions in the ocean model, interannual variability is minimal in ocean fluxes and is caused entirely by small interannual variations in anthropogenic emissions. Interannual variability in modeled terrestrial fluxes is larger and is caused primarily by interannual variations in temperature that affect respiration through the Q₁₀ mechanism. These variations correlate moderately well with CO₂ growth rate anomalies, but are too small in magnitude by about a factor of 3 to explain them entirely.

Between 1850 and 2004, the estimated cumulative ocean sink is 120 GtC with a 95% confidence range between 82 and 160 GtC. The average modeled ocean sink for the 1980s is 1.9 GtC yr⁻¹, and the average ocean sink for the 1990s is 2.2 GtC yr⁻¹. These values are somewhat higher than the estimates by McNeil et al. (2003), although the McNeil values are well within the 95% confidence interval. The cumulative oceanic carbon uptake from 1850-1994 is 95.2 GtC, which is roughly one standard deviation
smaller than the estimate by Sabine et al. (2004). Given the model structure, the two observational constraints on ocean carbon uptake conflict with each other. The cumulative flux estimate (Sabine et al., 2004) is above the maximum likelihood model fit while the decadal-scale flux estimate (McNeil et al., 2003). It is possible that adding a more realistic description of oceanic physics and carbon uptake (e.g., changes in oceanic deep water formation rates) would act to resolve this discrepancy (Bryden et al., 2005; Sarmiento, 1998).

2.3.5 Probabilistic predictions

This model and the associated parameter estimates can be used to make future predictions, and more importantly, to demonstrate the uncertainty of these predictions. We demonstrate this predictive uncertainty using the IPCC S550 emissions scenario, in which CO$_2$ concentrations are stabilized at 550 ppm (roughly twice that of preindustrial levels) (Fig. 2-7). Because concentrations are specified in the S550 scenario, allowable emissions are known perfectly if the strength of the carbon sink is known. Therefore, the uncertainty in allowable emissions is equivalent to the uncertainty in the carbon sink. The allowable emissions in 2005 are 8.3 GtC with a 95% confidence interval of 1GtC. The strength of the total sink in the same year is about 4.9 GtC with the same uncertainty range. The strength of the ocean sink in 2005 is about 2.2 GtC with a 95% uncertainty range of 1.4 GtC, and the strength of the terrestrial sink in 2005 is 2.6 GtC with a 95% uncertainty range of 1.8 GtC, demonstrating that the sum of the fluxes is better constrained than the individual fluxes. Given this emissions scenario, allowable
emissions peak around 2050, then drop as CO₂ concentrations begin to level off. The total sink strength peaks around 2065, after which the oceanic sink levels off and the effects of the respiration-temperature feedback begins to overtake effects of carbon fertilization. Uncertainty of all variables except oceanic flux increases with time; by the year 2100, the 95% uncertainty range of terrestrial flux is twice that of oceanic flux.

Figure 2-7: Predicted allowable emissions under the S550 stabilization scenario (a). The maximum likelihood (solid line) and 95% confidence intervals (dashed lines) of model predictions under an S550 stabilization scenario from the years 2005-2100 for ocean + terrestrial carbon uptake (b), terrestrial carbon uptake (c) and ocean carbon uptake (d). In these model simulations, CO₂ concentrations are prescribed and the allowable CO₂ emissions are calculated from the atmospheric balance of CO₂.
Which observation (given the same observation error) would constrain future allowable CO₂ emissions the most? We address this question by adding additional hypothetical observations that are assimilated into our model. The oceanic flux observation reduces the 95% confidence interval by about 10%. Terrestrial flux observations reduce this range by about 15%, or 50% more than oceanic flux observations (Fig. 2-8). The main reasons why this occurs are twofold: First, oceanic fluxes are already constrained by several observations in our analysis, whereas terrestrial fluxes are only constrained indirectly. As mentioned above, the uncertainty in the strength of the terrestrial sink is larger than the oceanic sink, especially in the latter part of the 21st century. Because most of the total sink uncertainty is related to the terrestrial parameters, reducing the uncertainty of these parameters has more power to reduce total sink uncertainty. Second, modeled terrestrial fluxes display interannual variability; thus, an observation of terrestrial flux combined with an observation of globally averaged temperature constrains the terrestrial parameters more than an ocean observation, which does not display interannual variability and only contains information about the long-term trend. According to most global inversion studies and more complex ocean models, oceanic fluxes show interannual variability around 0.5 GtC yr⁻¹ (McKinley et al., 2004). This oceanic flux variability is smaller than the estimated interannual variability of the terrestrial net carbon fluxes (Le Quere et al., 2003; Rodenbeck et al., 2003b).
Our model likely underestimates the utility of ocean flux observations by neglecting ocean flux variability. Yet, independent estimates of terrestrial fluxes are crucial to reduce sink uncertainty. Several observation systems could be utilized in order to produce global estimates of net or gross terrestrial fluxes. Satellite data, in combination with simple models and reanalyzed climate data, have been used to estimate annual global NPP (Nemani et al., 2003). Similarly, soil respiration chamber data have been scaled up to global estimates using satellite and climate data (Raich et al., 2002; Reichstein et al., 2003). FLUXNET, a network of over 200 eddy covariance towers, provides continuous direct measurements of net terrestrial fluxes (Baldocchi et al., 2001).

Figure 2-8: 95% confidence range of the allowable CO₂ emissions for the S550 stabilization scenario from the years 2005-2100. The solid black line is an estimate based on the currently available historical constraints of CO₂ concentrations and oceanic uptake. The dashed and dotted lines are estimates with the historical constraints plus an added hypothetical 2004 terrestrial flux constraint or an added hypothetical 2004 ocean flux constraint, respectively. Uncertainties of the hypothetical terrestrial and ocean flux constraints are both 0.4 PgC, which is equal to the McNeil et al. constraint uncertainty.
Although these measurements have small flux footprints, or representative areas, they respond coherently to large-scale climate anomalies (Ciais et al., 2005). The potential exists to upscale these measurements in order to produce a global terrestrial flux estimate, although such an estimate would have to consider uncertainties related to both measurement and representativeness errors.

2.4 Conclusion

We use a Bayesian synthesis inversion of the global carbon cycle observations to estimate past and predict future CO₂ sinks. The resulting parameter PDFs are not normally distributed, and significant correlations are present between parameters. Neglecting autocorrelation in the annual Mauna Loa CO₂ causes significant biases and overconfidence in model parameters and predictions. Neglecting interannual variability in global temperatures results in less constrained carbon cycle parameters. CO₂ growth rate anomalies combined with interannual temperature variations constrain the respiration temperature sensitivity (Q₁₀) parameter, which in turn constrains the carbon fertilization parameter (β).

The uncertainty in predicted allowable CO₂ emissions to follow a specific CO₂ stabilization path increases with time. Given the adopted model structure, the terrestrial sink is more uncertain than the oceanic sink given our model structure. Assimilating hypothetical observations of annual terrestrial and oceanic CO₂ fluxes show that terrestrial fluxes have more power to reduce sink uncertainty than observations of oceanic fluxes. Regardless of the relative importance of oceanic and terrestrial observations, this
study stresses the need for independent estimates of global-scale terrestrial fluxes. Continued monitoring of the carbon cycle will enable the design of better carbon cycle models and will assist in the design of possible carbon management strategies by reducing uncertainty about the strength of the future carbon sink.
Chapter 3

Causes of interannual variability in ecosystem-atmosphere carbon dioxide exchange in a northern Wisconsin forest using a Bayesian synthesis inversion

3.1 Introduction

Atmospheric concentrations of CO₂, as indicated by Mauna Loa Observatory, have risen about 20% since 1960 (Keeling and Whorf, 2005). The level of these concentrations in the future is likely a primary factor in the degree of climate change and resulting impact that occurs in the 21st century and beyond. Estimating anthropogenic CO₂ emissions is insufficient to predict CO₂ concentrations because roughly half of anthropogenic CO₂ emissions are currently absorbed by the terrestrial biosphere and the oceans. This sink magnitude is determined by the difference between the global CO₂ increase and the sum of fossil fuel emissions (Marland et al., 2003) and land use emissions (Houghton, 2003). Unlike anthropogenic CO₂ emissions, this sink of CO₂ has a great deal of interannual variability, with a 5 GtC yr⁻¹ range in magnitude over the 1990s. Although atmospheric inversion models indicate that the terrestrial sink is smaller than or equal in magnitude to the oceanic sink (Gurney et al., 2002; Rödenbeck et al., 2003a), as much as two thirds of total sink variability has been attributed to the terrestrial...
biosphere because the strength of the oceanic sink is probably more consistent on an interannual timescale (Bousquet et al., 2000; Rödenbeck et al., 2003b).

The net terrestrial sink is driven by two much larger gross ecosystem fluxes: photosynthesis and respiration. Climate, mainly through temperature, radiation and precipitation, has been shown to be a primary driver of interannual variability in both net primary productivity (Nemani et al., 2003) and respiration (Raich et al., 2002) in many ecosystems. Regional terrestrial CO$_2$ fluxes are also significantly correlated with climate indices such as the Southern Oscillation and Arctic Oscillation (Potter et al., 2003). Coupled carbon cycle/climate models predict global reductions in the strength of the terrestrial sink to varying degrees (Cox et al., 2000; Dufresne et al., 2002; Zeng et al., 2004), primarily due to a positive feedback between respiration and temperature. Although climate is an important driver of the terrestrial sink, other factors such as carbon fertilization and nutrient availability cannot be neglected (Cao and Woodward, 1998; Cramer et al., 2001). A better understanding of climate impacts on interannual variability in terrestrial fluxes and how this is linked with other factors will lead to better predictions of terrestrial sink strength.

A number of eddy covariance sites in FLUXNET (Baldocchi et al., 2001) have been operating for at least five years and have reported large ranges of interannual variability in annual net ecosystem exchange of CO$_2$ (NEE) (e.g. Barford et al., 2001; Carrara et al., 2003; Goulden et al., 1996; Hollinger et al., 2004). Determining the spatial representativeness of eddy covariance fluxes is an active research question. The typical footprint of an eddy covariance flux measurement is on the order of 1 km$^2$, and it has been shown that seasonal to annual sums of net ecosystem exchange can vary over
relatively small distances within areas that have a similar climate. Sums of NEE are not coherent across the Chequamegon Ecosystem-Atmosphere Study (ChEAS) region, which is a focused area for land-atmosphere interaction research in northern Wisconsin containing several eddy covariance flux towers. This lack of spatial coherence may be caused by differences in land cover type, soil type and stand age, among other factors (Desai et al., this issue). However, interannual variability of NEE may be spatially coherent even if seasonal and annual sums are not. Different ecosystems may have similar responses to climate anomalies, causing spatially coherent patterns of flux variability on spatial scales similar to that of climate variability. For example, several European eddy covariance sites had a similar NEE response to the anomalously warm and dry summer of 2003 (Ciais et al., 2005). Thus, climate-driven models of NEE variability derived from a single site may be representative of areas larger than the tower footprint.

We present and analyze 7 years of eddy covariance data from the WLEF tower, which is part of both the AmeriFlux Network and the Chequamegon Ecosystem-Atmosphere Study (ChEAS). WLEF is a 447 m TV transmitter which has been instrumented to collect CO₂ and energy fluxes at multiple heights. The footprint, or area represented by the flux measurement, is up to 100 times as large in area as typical surface layer towers. Unlike most FLUXNET sites, WLEF has a footprint that encompasses several vegetation types including upland deciduous and coniferous forests, recently logged areas, and wetlands. WLEF has multiple flux measurement levels that can be synthesized to create an NEE product that has a more constant flux footprint from day to night than NEE products from a single level (Davis et al., 2003). WLEF is also part of
the National Oceanographic and Atmospheric Administration’s tall tower CO₂ monitoring network, with both flask and continuous high-precision CO₂ measurements (Bakwin et al., 1998).

The objective of this study is to develop a simple climate-driven model that explains seasonal and interannual variability in NEE at the WLEF site. Probability density functions (PDFs) of model parameters are estimated using the Markov Chain Monte Carlo (MCMC) framework. Modeled seasonal and annual NEE sums are compared to sums derived from observations that have been gap-filled using consistent methods (Desai et al., 2005; Falge et al., 2001). Random and systematic errors in the fluxes are evaluated and compared to model uncertainties. This study tests three main hypotheses:

1. The range of seasonal and interannual variability in NEE of CO₂ observed at WLEF is statistically significant compared to random uncertainty.

2. At both seasonal and annual timescales, correlations between soil moisture anomalies and NEE anomalies are statistically significant. Similarly, correlations between soil temperature anomalies and NEE anomalies are statistically significant.

3. A simple climate-driven model calibrated with observations produces statistically significant correlations with observed patterns of interannual variations in season-averaged and annual NEE from 1997-2004 at WLEF.
3.2 Methods

3.2.1 Data sources

WLEF is located in the Park Falls Ranger District of the Chequamegon National Forest, about 15 km east of Park Falls in northern Wisconsin (45.9459° N, 90.2723° W). The surrounding forest, with a maximum canopy height of 25 m, consists of upland hardwood forest and lowland wetlands. The elevation in this region varies between about 470 m and 500 m; saturated soils and wetlands are typically found in the lower elevations. The landscape is described in further detail by Davis et al. (2003) and Mackay et al. (2002).

Fluxes of water vapor, virtual temperature, CO₂ and momentum were measured at three levels: 30, 122 and 396 m. Sonic anemometers (Campbell Scientific, Inc., Logan, UT, model CSAT3 and Applied Technologies, Inc., Longmont, CO, 3D K-type) were used to measure high-frequency variations in three components of wind and virtual temperature, and closed path infrared gas analyzers (IRGAs) (Li-Cor, Lincoln, NE, model 6262) were used to measure high-frequency (5Hz) variations in concentrations of CO₂ and water vapor. The IRGAs were calibrated with high-precision, high-accuracy CO₂ mixing ratio measurements maintained by NOAA-CMDL (Bakwin et al., 1998; Zhao et al., 1997). The flux measurement methodology is described in greater detail by Berger et al. (2001). This analysis does not include any flux observations from the year 2002, when data are suspect due to problems with the data acquisition system.

Meteorological data are collected at the WLEF tower and at several nearby stations, including air temperature ($T_a$), 5 cm soil temperature ($T_s$),
photosynthetically active radiation (PAR), and volumetric soil water content (SWC). \( T_a \) and relative humidity are measured at 30, 122 and 396 m on the tower (Vaisala, Helsinki, Finland, model HMP45C). PAR is measured by LiCor quantum sensors (Li-Cor, Lincoln, NE, model LI-190SZ). SWC is not measured at WLEF, but is measured at the Mixed Upland site and at the Willow Creek flux tower at 10 cm depth by reflectometer probes (Campbell Scientific, Inc., Logan, UT, model CSI CS615). Bad meteorological data caused by instrument failure or shading of the PAR sensor by the tower are filled with data from the nearby Willow Springs or Mixed Upland sites, or from the Willow Creek flux tower (Cook et al., 2004). If data from all sites are bad, gaps are filled with monthly diurnal mean values.

### 3.2.2 Error analysis

In any study of interannual variability at an eddy covariance site, it is critical to show that the observed variability is not caused by biases, noise, or gap-filling techniques. At WLEF, there are three key sources of random error that confound estimates of interannual variability: 1) gap-filling uncertainty related to the amount of missing data, choice of gap-filling model, and goodness of model fits; 2) turbulent sampling error, which causes large uncertainties in the hourly measurements and may be significant at seasonal to annual timescales; and 3) uncertainty caused by interannual variations in the distribution of levels that is used to compute the “preferred” NEE product as described by Davis et al. (2003). Even with an identical distribution of levels in the preferred NEE product from year to year, some interannual variability may exist in
the size, orientation of, and distribution of vegetation types in the average tower footprint; however, these effects are not analyzed in this study due to the lack of an accepted description of flux footprints for measurement altitudes above the surface layer. Simple analyses of fluxes as a function of wind direction suggest that this is a second-order effect. Examining flux footprints in the mixed layer may enable this effect to be evaluated quantitatively in future studies.

A large number of other potential systematic error sources exist, such as errors in calibration, the selection of a $u^*$ threshold or low- and high-frequency spectral attenuation. These systematic errors depend on instrumentation and data processing, and may vary among sites (Massman and Lee, 2002). For example, the averaging time of one hour may be too short under some conditions at 122 m and 396 m, which may cause attenuation of low-frequency components of the flux (Lenschow et al., 1994; Sakai et al., 2001). Although these sources of error can cause absolute values of NEE sums to change significantly, they are not expected to have significant impacts on interannual variability because they generally affect measurements similarly during all years. Data processing methodology, including calibration, correction of $CO_2$ fluxes for high frequency attenuation, and the rotation of sonic anemometers are described in detail by Berger et al. (2001).

### 3.2.2.1 Data screening

In general, annual NEE sums are a strong function of the $u^*$ threshold that is chosen to screen low turbulence nighttime data. Uncertainty in annual NEE in several
eddy-covariance tower studies has been estimated by integrating NEE measurements over a range of \( u^* \) thresholds (e.g. Hollinger et al., 2004). The choice of a \( u^* \) threshold generally does not affect the annual values relative to each other (Fig. 3-1), especially for values close to the chosen \( u^* \) threshold of 0.2 m s\(^{-1}\). Because we focus on interannual variability, the systematic uncertainty related to the choice of \( u^* \) threshold is not considered further. During the 1997-2004 period, the WLEF region was source of carbon to the atmosphere regardless of the choice of \( u^* \) threshold (Fig. 3-1). Consistent carbon sources are not common in the AmeriFlux and FLUXNET literature, but have been observed at eddy-covariance towers in the Amazon rain forest (Saleska et al., 2003) and in a European mixed forest (Carrara et al., 2003).

Figure 3-1: Cumulative annual NEE (a) and percentage of hours gap-filled (b) as functions of the \( u^* \) threshold that is used to screen nighttime data. The dashed vertical line indicates the value of the \( u^* \) threshold that was used in this paper: 0.2 m s\(^{-1}\).
For this study, nighttime NEE were screened when the value of $u^*$ is lower than 0.2 m s$^{-1}$. Typically, $u^*$ thresholds are chosen to be the point above which mean nocturnal respiration is roughly constant as a function of $u^*$ (Goulden et al., 1996). At WLEF, nighttime NEE (normalized by $T_a$) continues to increase as a function of $u^*$ until values of $u^*$ reach 0.35 or 0.4 m s$^{-1}$. Data become sparse at such high values of the $u^*$, causing the gap-filling algorithm to become biased or fail. It is evident that estimates of annual NEE sums decrease as the $u^*$ threshold approaches 0.4, which occurs despite the increase in observed nighttime NEE (Fig. 3-1). In this case, lack of data forces non-growing season data to be used in order to produce functional fits to estimate growing season respiration. Nighttime observations at WLEF indicate that non-growing season respiration is generally smaller than growing season respiration given the same temperature. Average values of $u^*$ are also lower during the growing season: a choice of $u^*$ threshold = 0.4 m s$^{-1}$ would screen about 66% of nighttime non-growing season data (Jan-May; Sep-Dec) and more than 80% of growing season data (June-August), while a choice of $u^*$ threshold = 0.2 m s$^{-1}$ screens about 40% of nighttime non-growing-season data and about 55% of growing-season data. The $u^*$ threshold value of 0.2 m s$^{-1}$ is thus chosen to be consistent with Davis et al. (2003) and because a choice of the threshold within the range 0 to 0.25 m s$^{-1}$ provides sufficient data for the gap-filling algorithm while having little effect on NEE variations between years.
3.2.2.2 Gap-filling methodology

The gap-filling technique is based on the nonlinear regression method as described by Falge et al. (2001). Each missing hour in the data is filled by the modeled sum of ecosystem respiration ($R_E$) and gross ecosystem production (GEP). $R_E$ is modeled using a modification by Law et al. (2002) of the function originally described by Lloyd and Taylor (1994):

$$R_E = R_{10} \exp \left( \frac{E_a}{R} \left[ \frac{1}{T_{ref}} - \frac{1}{T_a} \right] \right), \quad 3.1$$

where $R_{10}$ is the base respiration at an $T_a$ of $10^\circ$C, $E_a$ is the activation energy in Joules mol$^{-1}$, $T_{ref} = 283.15$K, and $T_a$ is the air temperature in Kelvin. This function is also used by Noormets et al. (this issue) in a cross-comparison study of respiration at 14 eddy covariance sites in the ChEAS region. $T_a$ is used as the input variable rather than $T_s$ because nighttime NEE at WLEF is observed to vary as a function of $T_a$, even when $T_s$ is held constant. As a result, $R^2$ values for fits using $T_a$ are higher than those for fits using $T_s$. This increased correlation when using $T_a$ was also noted in a multi-site EUROFLUX analysis (Van Dijk and Dolman, 2004). The $E_a$ and $R_{10}$ parameters are determined by fitting the model to $u^*$-screened nighttime NEE. Nighttime is defined as when observed PAR is equal to zero; no photosynthesis occurs during these hours. Fit parameters are obtained for every day of each of the seven years by fitting all points within a one-month moving window using a nonlinear least-squares regression routine. This method of parameter optimization assumes that all hourly data points have the same uncertainty, although in reality flux uncertainty is a function of boundary layer depth, stability and
other variables (Berger et al., 2001; Lenschow and Stankov, 1986). If too few points exist for a good fit (< 100 points), the one-month fitting window is expanded one day at a time until this requirement is met. For gaps greater than 15 days (360 points), NEE is filled with the 7-year mean monthly diurnal cycle.

“Observed” gross ecosystem productivity (GEP) is calculated by subtracting modeled respiration, which is computed from Eq. 3.1 using daytime T_a, from daytime observed NEE. GEP is then modeled using a Michaelis-Menten reaction rate equation:

\[
GEP = \frac{GEP_{\text{max}} \times PAR}{PAR_{1/2} + PAR},
\]

where \(GEP_{\text{max}}\) and \(PAR_{1/2}\) are the model parameters. As with respiration fits, the \(GEP_{\text{max}}\) and \(PAR_{1/2}\) parameters are obtained for every day using the one-month moving window, which can be expanded if not enough points exist for a good fit.

The amount of data missing before u* screening ranges from roughly 10-20%, and an additional 20-25% of data are screened using a u* threshold of 0.2 m s^-1 (Fig. 3-1b). Missing data most often result from failure of the sonic anemometers due to harsh conditions on the WLEF tower, especially at the higher levels. Missing data also result from power outages, computer malfunctions, pump failures and temporary sonic outages due to condensation and precipitation. Suspect high-frequency data are also screened as part of data processing, resulting in additional missing fluxes. Depending on the nature of the problem, data gaps may range in size from a single hour to several weeks.
3.2.2.3 Gap-filling uncertainty

Overall, gap-filling introduces a small amount of random error on annual estimates of NEE (Falge et al., 2001). If the gap fraction is large or if the gap-filling model is biased, gap-filling may also introduce systematic error to NEE sums. The reduction in NEE sum at large values of the $u^*$ cutoff point is caused by a bias in the gap-filling method when the fitting window is large (Fig. 3-1). We assume that using a $u^*$ cutoff of 0.2 m s$^{-1}$ results in unbiased gap-filling and do not consider this possible systematic error. To quantify the random uncertainty, a Monte Carlo experiment was performed using a method similar to Griffis et al. (2003) and Desai et al. (2006). “Perfect” NEE datasets were created for each year 1997-2004 by gap-filling the observed NEE and imposing simulated random turbulent sampling error on the filled hours.

The effects of gap-filling were then analyzed for each year using the following steps: 1) 100 identical “perfect” NEE datasets were created; 2) A different set of data gaps was created for each dataset, but all datasets contained the same fraction of missing data; 3) The gap-filling algorithm was applied to each dataset; 4) The resulting range of annual NEE values over the 100 datasets was calculated. Gap locations were created using a random number generator. In each of these 100 datasets, the percentages of both missing and screened data were equal to that in the original observed NEE time series. Missing data gaps were treated differently than gaps resulting from $u^*$ screening: Missing data gaps were removed in random blocks (of up to 120 consecutive data points), while screened data were removed randomly as single hours from the remaining nighttime data.
3.2.2.4 Turbulent sampling uncertainty

Hourly eddy-covariance measurements of NEE contain random uncertainties related to turbulent sampling that may be as large as the measured fluxes themselves at WLEF (Berger et al., 2001). These errors are functions of flux magnitude, stability, boundary layer depth and measurement height (Lenschow and Stankov, 1986). Because turbulent sampling errors from different hours are independent of each other, integrated sums of NEE over seasonal to annual timescales generally have much smaller errors. However, these errors may still be large enough to cloud the general picture of seasonal or interannual variability. Davis et al. (2003) estimated the random error in the 1997 NEE sum empirically by calculating the standard error of daily cumulative NEE measurements. This estimate, which includes turbulent sampling error, random instrumental noise and variations due to weather events, was 19 gC m\(^{-2}\).

In this study, hourly turbulent sampling uncertainty is estimated as the standard deviation of differences between observed and modeled NEE. Modeled NEE is the sum of fit equations (1) and (2) using hourly PAR and \(T_a\) as input. Turbulent sampling uncertainty is assumed to vary as a function of time of day and season. Estimates will include random instrumental error as well (e.g. electronic noise in the wind and CO\(_2\) measurements), but this is typically a small contribution since noise from the two instruments must be correlated to influence the eddy-covariance calculation. Turbulent sampling uncertainties were calculated for every hour of the day within four periods: December-March, April-May, June-August and September-November (Fig, 3-2). Errors are generally larger in the growing season when flux magnitudes are larger, and in the
daytime when eddies are larger. These errors are assumed to be independent; therefore, annual error estimates are calculated by taking the square root of the sum of hourly errors. Estimates of error for annual sums are smaller than that reported by Davis et al. (2003); their value was obtained by computing a standard deviation of all hourly flux values, and thus includes diurnal, synoptic and seasonal variability.

Figure 3-2: Diurnal cycle of estimated hourly random error (turbulent sampling plus instrumental noise) in NEE measurements in December-March (a), April-May (b), June-August (c) and September-November (d). Hourly random error was estimated as the standard deviation of the difference between modeled and observed NEE over each season. Modeled values of NEE were calculated using gap-filling regression parameters.
We also assume that the nonlinear regressions derived for gap filling produce unbiased estimates of NEE; thus, our estimates of turbulent sampling errors are likely overestimates because errors in model structure are interpreted as turbulent variability. The method of weighting points in the gap-filling regressions does not affect our estimates of turbulent variability. Richardson et al. (2006) developed a method to independently assess turbulent sampling errors by comparing fluxes on adjacent days under similar conditions; such a procedure may be more accurate and more statistically sound than the method used above. Because turbulent sampling errors and instrumental noise are both independent, estimated random uncertainties of seasonal and annual sums may be computed by taking the square root of the sum of the squared errors of all of the hourly fluxes within the period.

3.2.2.5 Level selection uncertainty

The “preferred NEE” algorithm developed by Davis et al. (2003) selects the appropriate level to use in the final NEE product based on micrometeorological conditions and data availability. Generally, NEE is selected from the 30-meter level at night and is an average of the 122 m and 396 m level during the daytime. Nighttime data from the upper two levels are not used because these levels are generally above the stable boundary layer, where storage fluxes are the primary component of NEE. Storage fluxes are calculated from mixing ratio data and likely have a different footprint than turbulent flux measurements (Gloor et al., 2001). During the morning transition, 30 m NEE is used to avoid the large uptake signal observed at 122 and 396 m. During this time, horizontal
advection is suspected to be significant, causing the higher-level measurements to be biased (Yi et al., 2000). Effects at the lowest level are small because advection influences the NEE estimate by changing column storage, and a shallow column minimizes storage. When data from the desired level are missing, the preferred NEE algorithm selects data from any available level. Differences in NEE between levels coupled with interannual variations in data availability can cause error that is misinterpreted as variability in seasonal or annual NEE sums.

This possible bias was examined by constructing and analyzing a complete hourly NEE dataset for the three tower levels. This hourly datasets consisted of monthly diurnal mean values, averaged over times only during which all three levels had good data. We applied the preferred NEE algorithm to this complete dataset to obtain a value of annual NEE using the optimal distribution of levels. Then for each of the seven study years, the actual distribution of levels chosen by the preferred NEE algorithm on the actual dataset was used on the complete dataset to obtain annual NEE sums. The difference in NEE when the optimal distribution of levels is used versus when the actual distribution of levels is used is an estimate of the error caused by interannual variations in data availability. The range of this error was about 20 gC m$^{-2}$ when using the preferred NEE algorithm as described by Davis et al. (2003).

The average diurnal cycle of NEE varies considerably as a function of season and measurement level (Fig. 3-3). During the nighttime, there is relatively close agreement among levels, even though the upper two levels are generally outside the stable boundary layer and storage flux is the primary component of NEE. Between 5 and 10 LST, the morning transition problem is evident at 122 and 396 m in the form of anomalously low
NEE values. During the daytime in the spring and summer, NEE values diverge. Uptake at 30 m is considerably lower than that at 122 meters. The grassy clearing around the WLEF tower may comprise a significant fraction of the flux footprint at the 30-meter level during the day. Fluxes from this clearing are not representative of the surrounding mixed forest ecosystem (Wang et al., 2006). Given the same meteorological conditions, using more 30-meter daytime growing season data in one year relative to another will produce a false signal of interannual variability with an estimate of lower carbon uptake.

Uptake at 396 m is considerably greater than at 122 m; using more 396-meter daytime growing season data in one year relative to another would also produce a false interannual variability signal with estimates of increased uptake. The reason for the difference in daytime NEE between 122 m and 396 m is unknown at this time. The distribution of vegetation types in the 396 m footprint is similar to that in the 122 m footprint, and footprint analysis demonstrates that grassy clearing has a minimal effect on both of these higher levels (Wang et al., 2006). Estimates of regional CO₂ fluxes from simple regional inversion techniques (Bakwin et al., 2004; Helliker et al., 2004) and a regional upscaling of multiple stand-level towers (Desai et al., this issue) tend to agree more closely at 396 m than at 122 m, indicating that measurements from this level may be more representative of the surrounding region. Unfortunately, the 396 m NEE data are sparse because of frequent sonic anemometer failure.

The preferred NEE algorithm was therefore modified to default to the 122 m in the daytime; if this level is missing, an average of 30 and 396 m NEE is used. This modification limits possible biases that may occur by using only 30 m or 396 m data and reduces the bias caused by interannual variations in data availability by about a factor of
two (10 gC m$^{-2}$). This estimated bias is added as a correction to all reported NEE sums.

All further analysis uses this modified version of the preferred NEE algorithm.

---

Figure 3-3: Diurnal cycle of NEE as measured at 30, 122 and 396 meters during winter (a), spring (b), summer (c), and autumn (d). Data are from hours during which measurements exist for all three levels from 1997 through 2004.
### 3.2.2.6 Total random uncertainty

The total uncertainty results from random errors including turbulent sampling error and gap-filling uncertainties. Estimates of total uncertainty are produced for each year by adding random, normally distributed turbulent sampling errors with the previously estimated standard deviations (Fig. 3-2) to the observed values in each of the 100 datasets derived using the Monte Carlo gap-filling uncertainty estimation method. No simulated turbulent variability is added to gap-filled values; hence, the estimated turbulent sampling uncertainty of NEE sums decrease with increasing gap fraction. This decrease in turbulent sampling uncertainty is compensated by increased gap-filling uncertainty.

### 3.2.3 Interannual model

The objective of the interannual model is to reproduce variability in seasonal and annual sums of NEE with a single set of model parameters. The interannual model is similar to the gap-filling model in that hourly NEE are assimilated to optimize model parameters. Instead of estimating a set of model parameters for each day of every year, the interannual model is used to estimate a single set of parameters for the entire WLEF dataset. Because this model assimilates hourly observations of NEE, it must also accurately reproduce variability at shorter timescales. The gap-filling models for ecosystem respiration (Eq. 3.1) and gross ecosystem productivity (Eq. 3.2) are extended to incorporate phenology and soil moisture effects and combined to produce a single model of NEE.
3.2.3.1 Respiration submodel

Respiration on diurnal to synoptic timescales is assumed to follow an exponential relationship (Lloyd and Taylor, 1994), which is also used in the gap-filling model (Eq. 3.1). To accommodate changes in respiration over seasonal timescales, this equation is extended so that the base respiration term incorporates the effects of SWC and $T_s$ (as a proxy for phenology). A one-month running mean filter is used to smooth $T_s$ data in order to average over several synoptic cycles. We use a single set of model parameters to predict respiration during the spring, summer and autumn months (defined as April-November). Respiration in December through March is not modeled because interannual variability of NEE during the winter months is minimal, and an appropriate model during this time period would require consideration of the effects of frozen soil and snow cover (Brooks et al., 2005).

Our formulation of the interannual $R_E$ submodel is based on both current mechanistic understanding of ecosystem respiration and the behavior of daily parameters from the gap-filling model. Results of a recent EUROFLUX study indicate that base respiration varies significantly at many sites over the course of a year while the temperature sensitivity $E_a$ remains relatively constant (Van Dijk and Dolman, 2004). If $E_a$ is assumed to be constant, the daily base respiration parameter $R_{10}$ derived from gap-filling has a strong dependence on $T_s$ (Fig. 3-4a). After the linear dependence on $T_s$ is removed, the residual $R_{10}$ is a weak function of SWC when $E_a$ is held constant (Fig. 3-4b). There appears to be an optimal value of SWC at which respiration peaks, then falls off above and below this value. We approximate this by introducing an optimal SWC
parameter, and a factor to describe the rate of reduction as a function of distance from the optimal SWC.

A similar behavior was noted in a chamber flux study in the ChEAS area that analyzed differences in respiration along soil moisture gradients (Martin and Bolstad, 2005). Skopp et al. (1990) also predicted that an optimal SWC should exist, above or below which soil microbial activity declines. Reichstein et al. (2003) recently derived a

Figure 3-4: Bin-averaged dependence of daily $R_{10}$ and $GEP_{\text{max}}$ parameters on soil temperature ($T_s$) and soil moisture (SWC) when assuming the air temperature sensitivity $E_a$ and half saturation constant PAR$_{1/2}$ are constant. The $T_s$ dependence of $R_{10}$ is described well by a linear fit represented by the solid line (a). The residuals of this fit are a function of SWC (b). Similarly, the $T_s$ dependence of $GEP_{\text{max}}$ is described well by a linear fit represented by the solid line (c), and the residuals are a function of SWC (d).

A similar behavior was noted in a chamber flux study in the ChEAS area that analyzed differences in respiration along soil moisture gradients (Martin and Bolstad, 2005). Skopp et al. (1990) also predicted that an optimal SWC should exist, above or below which soil microbial activity declines. Reichstein et al. (2003) recently derived a
similar respiration model that includes SWC to explain variability in chamber measurements across 17 sites. In their model, base respiration is related to SWC through a half-saturation relationship. Such a formulation was found to be inadequate at WLEF because of the observed dropoff in respiration during very moist conditions; this was not problematic in Reichstein et al. (2003) because the sites analyzed were generally Mediterranean in climate and not as moist as northern Wisconsin. Large pulses of respiration have also been observed at dry eddy covariance sites following rainfall events (Xu et al., 2004). Because WLEF is generally a moist site, we do not attempt to incorporate this effect in our model.

A linear dependence of base respiration on $T_s$ is used in the model to estimate base respiration during the transitional seasons of spring and autumn. Because $R_E$ includes both heterotrophic and some components of autotrophic respiration, base respiration is expected to increase during leafout and decrease during senescence. Phenologists often use cumulative sums of growing degree days (GDD) to model leafout; one example is the formulation in the PnET-Day model (Aber et al., 1996). This study aims to find a single driving variable that predicts both leaf-out and senescence reasonably well in order to limit the number of model parameters. We note that the $R_{10}$ dependence on $T_s$ that is approximated well with a linear fit (Figure 3-4a) and is reasonably consistent from spring to autumn. $T_s$ also has a high correlation with GDD in the springtime. The linear fit of $R_{10}$ to $T_s$ has a nonzero y-intercept because small positive values of NEE are observed when soils are frozen or snow-covered at WLEF. We account for this effect in the respiration submodel, but do not attempt to predict $R_E$ at values of $T_s$ under 2°C.
In this model, the activation energy $E_a$ is a single parameter and therefore assumed to be constant throughout the year at WLEF. Effectively there are two temperature sensitivities: short-term and long-term sensitivities are represented by the $E_a$ and linear dependence of $R_{10}$ parameters respectively. Separating the timescales respiration temperature dependence is supported by Reichstein et al. (2005). $E_a$ may also depend on SWC (Reichstein et al., 2003). Although a small increase in $E_a$ does occur with increasing SWC at WLEF, Noormets et al. (this issue) observed that respiration equations incorporating a SWC dependence of $E_a$ fail to converge or produce unstable parameter estimates for sites within the ChEAS region.

Given the above observations and assumptions, the submodel is:

$$R_E = R_{10}(T_s) \cdot f_R(SWC) \cdot \exp \left[ \frac{E_a}{R} \left( \frac{1}{T_{ref}} - \frac{1}{T_a} \right) \right],$$

where

$$f_R(SWC) = 1 - a_R \left( \frac{SWC - SWC_{optR}}{SWC_{optR}} \right).$$

$$R_{10}(T_s) = R_{10\text{froz}} + R_{10\text{fac}} \cdot T_s$$

This submodel contains five varying parameters ($E_a, a_R, SWC_{optR}, R_{10\text{froz}}, R_{10\text{fac}}$). The base respiration $R_{10}$ ($R_E$ at a $T_a$ of 10°C) varies as a function of $T_s$. $f_R(SWC)$ is a function that acts to reduce this base respiration under non-ideal SWC conditions. $E_a$ (Joules) is the activation energy, which controls the sensitivity of respiration to $T_a$. $R_{10\text{froz}}$ ($\mu$mol m$^{-2}$ s$^{-1}$) is the value of base respiration at $T_a = 10^\circ$C when $T_s = 0^\circ$C. $a_R$ (dimensionless) is a linear factor that determines the strength of SWC dependence. $SWC_{optR}$ (volumetric soil water fraction) is the optimal value of SWC at which $f_R(SWC)$
is equal to one. Deviations from this optimal value in either direction cause a reduction in base respiration.

### 3.2.3.2 Photosynthesis submodel

Hourly to synoptic GEP is assumed to follow the Michalis-Menton relationship (Eq. 3.2). As with the interannual respiration submodel, this equation is then extended so that the $GEP_{\text{max}}$ term incorporates the effects of SWC and $T_s$ (as a proxy for phenology) to predict GEP during the spring, summer and autumn months (defined as April-November).

We assume that the phenological and moisture dependence of GEP is manifested through the $GEP_{\text{max}}$ parameter and not the half saturation constant $PAR_{1/2}$. Assuming a constant $PAR_{1/2}$, the $T_s$ dependence of the daily gap-filling fit parameter $GEP_{\text{max}}$ in the growing season is linear because of the high correlation between $T_s$ and GDD during the spring. This dependence is also similar to that of $R_{10}$ (Figure 3-4c). After the linear dependence on $T_s$ is removed, the residual $GEP_{\text{max}}$ dependence on SWC is similar to that of the residual $R_{10}$ with an optimal value of SWC. The reduction of GEP at low SWC is likely to be caused by increased stress associated with drought conditions and high values of VPD. SWC is also negatively correlated with time of year during the growing season, and increased leaf age has also been shown to be a factor in reducing GEP (Wilson et al., 2001). The mechanism for reduced GEP at high values of SWC is unclear at this time. Thus, we have made $f_{GEP}(SWC)$ identical in structure to $f_R(SWC)$ for simplicity.
Given the behavior of the gap-filling model parameters and mechanistic constraints, we adopt the following form for the interannual GEP model:

\[
GEP = GEP_{\text{max}}(T_s) f_{\text{GEP}}(\text{SWC}) \frac{\text{PAR}}{\text{PAR}_{1/2} + \text{PAR}}, \text{ where}
\]

\[
f_{\text{GEP}}(\text{SWC}) = 1 - a_{\text{GEP}} \left( \frac{(\text{SWC} - \text{SWC}_{\text{optGEP}})}{\text{SWC}_{\text{optGEP}}} \right)
\]

\[
GEP_{\text{max}}(T_s) = GEP_{\text{fac}} \cdot T_s
\]

The GEP model contains four varying parameters (\text{PAR}_{1/2}, a_{\text{GEP}}, \text{SWC}_{\text{optGEP}} and \text{GEP}_{\text{fac}}). GEP_{\text{max}}(T_s) is the value of GEP under saturated PAR and ideal conditions of SWC and is a function of T_s. Unlike the interannual respiration equation (Eq. 3.3), we have assumed that GEP = 0 if the soil is frozen or snow-covered (T_s = 0). GEP_{\text{fac}} (\mu\text{mol m}^{-2} \text{ s}^{-1} \text{ K}^{-1}) is a linear factor that controls the dependence of GEP_{\text{max}} on T_s, similar to the \text{R}_{\text{fac}} parameter in Eq. 3.3.

As in the respiration submodel, \(f_{\text{GEP}}(\text{SWC})\) is a function that limits GEP_{\text{max}} in non-ideal soil moisture conditions. Although a dependence on VPD rather than SWC may be more mechanistically sound, we choose the SWC formulation because it is exactly analogous to the respiration equation and therefore allows a direct comparison of the \(a_{\text{GEP}}\) and \(a_R\) parameters. This function is a fractional multiplier to GEP_{\text{max}}(T_s) and always range between zero and one. PAR_{1/2} (\mu\text{mol m}^{-2} \text{ s}^{-1}) is the value of PAR at which GEP is one half of GEP_{\text{max}} under ideal moisture conditions. \(a_{\text{GEP}}\) (dimensionless) and \(\text{SWC}_{\text{optGEP}}\) (volumetric fraction) are exactly analogous to \(a_R\) and \(\text{SWC}_{\text{optR}}\) as described in the respiration submodel.
3.2.4 Nonlinear optimization methodology

Markov Chain Monte Carlo (MCMC) is a powerful method to assimilate observations into nonlinear models that are computationally fast to evaluate and have a modest number of parameters. MCMC simulates direct draws from a joint probability density function, making no assumptions about the shape of this distribution. Prior information about parameters can be included using a Bayesian framework; using diffuse uniform priors recovers the maximum likelihood approach in a frequentist sense. We use the Metropolis-Hastings algorithm (Hastings, 1970; Metropolis et al., 1953), which has shown to be relatively robust (Qian et al., 2003).

The posterior probability $p_{post}$ of a sample from the joint parameter distribution $\theta_k$ is a function of the prior probability $p_{prior}$ and the observations $x$, and is calculated as follows:

$$p_{post}(\theta_k | x) = \frac{L(x | \theta_k) p_{prior}(\theta_k)}{\sum_{i=1}^{N} L(x | \theta_i) p_{prior}(\theta_i)}$$

The denominator on the right-hand side of this equation is a normalization factor such that the sum of posterior probabilities over all possible sets of parameters is equal to one. The construction of the Markov chain only requires knowledge of probability and likelihood ratios, so that we may ignore this normalization factor during this step. The function $L$ is the likelihood of observing all of the observations $x$ given the set of parameters $\theta_k$. If all observations are independently and identically distributed (IID) and follow a normal distribution, then the likelihood is:
The likelihood of a set of \( n \) observations in \( \mathbf{x} \) is the product over the likelihoods of individual observations \( x_1, x_2, \ldots, x_n \) at times \( t_1, t_2, \ldots, t_n \). We assume that each observation is sampled independently from a Gaussian distribution with a mean \( x_i \) and standard deviation \( \sigma_i \) representing the variability due to observation noise and internal variability that is not captured by the model. Observation error is estimated using the method described in 3.2.2.4 and differs in magnitude from one observation to the next.

This formulation of the likelihood function follows previous studies using MCMC that assume IID and normally distributed variability (Braswell et al., 2005; Hargreaves and Annan, 2002). There is some indication that random flux error may more closely follow a double exponential rather than a Gaussian distribution (Richardson et al., 2006), but we have chosen to implement MCMC in a Gaussian framework to be consistent with previous studies. The IID assumption is useful and reasonable approximation only if both observation errors process noise are IID. Observational error related to turbulent sampling is likely to be independent because turbulent eddies in the atmospheric boundary layer are short-lived (Lenschow and Stankov, 1986; Berger et al., 2001) compared to the hourly flux measurement interval. Process error may not be independent, resulting in temporally autocorrelated residuals between modeled and observed NEE and therefore violating the IID assumption of Eq. 3.6. We address the potential problem of autocorrelation by thinning observations of NEE such that the residuals do not display significantly significant autocorrelation. This was achieved by thinning the hourly observations by a factor of five.
Individual parameter PDFs are projections of the joint posterior PDF taken directly from the Markov Chain after it has converged. To examine the robustness of the model parameters, we run this assimilation three times: (1) base case with hourly observed NEE data, (2) test case with only nighttime NEE data, and (3) test case with only daytime NEE data. Hours that have been gap-filled are not used in the assimilation. Only model predictions from the base case (1) are used to compare modeled and observed NEE as a function of season and time of day. The test cases (2) and (3) are run to explore how model parameters vary between nighttime and daytime, especially the respiration parameters in equation (3). If the interannual model of respiration is a mechanistically correct description of this process and if the observations are unbiased, we would expect that the respiration parameters remain constant in all three cases. If there are important differences between nighttime and daytime respiration that are not captured by the climate relationships in the simple model, or if the observations are biased in an unknown way, we would expect the parameters to change between daytime and nighttime with base case parameters being an average of the two. Parameters for GEP are not estimated in the nighttime case (2), and any differences between GEP parameters in cases (1) and (3) are caused by differences in partitioning of model gross fluxes.

3.2.5 Selecting Bayesian Priors

In a Bayesian framework, choices of priors for model parameters affect the posterior parameters and uncertainty ranges. Prior information provides important
constraints on model parameters in other eddy covariance data assimilation studies (Braswell et al., 2005; Knorr and Kattge, 2005). Using uniformly distributed priors over an infinite range of values is equivalent to the frequentist maximum likelihood approach. In this study, we attempt to use as little prior information as possible except to exclude physically unreasonable parameter values. All prior distributions are uniform, but are chosen to be within physically reasonable ranges as dictated by the data and existing literature.

The $R_{10fac}$ parameter was assumed to be between 0 and 1 $\mu$mol m$^{-2}$ s$^{-1}$ K$^{-1}$. The upper bound implies a base respiration rate of 15 $\mu$mol m$^{-2}$ s$^{-1}$ at a typical growing season $T_s$ of 15$^\circ$C, which is more than twice the magnitude of the largest daily $R_{10}$ fit parameter. The analogous GEP$_{fac}$ parameter ranges between 0 and 4 $\mu$mol m$^{-2}$ s$^{-1}$K$^{-1}$. The upper bound implies a maximum GEP of 60 $\mu$mol m$^{-2}$ s$^{-1}$K$^{-1}$ at a $T_s$ of 15$^\circ$C, again more than twice the magnitude of the largest daily GEP$_{max}$ parameter. The SWC$_{opt}$ prior bounds for both the RE and GEP submodels are set to the range of observed SWC at WLEF; optimal SWC is assumed to lie in this range. Skopp et al. (1990) reported that the optimal value of SWC for microbial activity in several experiments across varying soil types was about 60% of water-holding capacity. The upper bound for the factors $a_R$ and $a_{GEP}$ is set to 1; values larger than this would cause negative values of RE or GEP during extremely dry or moist conditions. The prior range of the activation energy $E_a$ corresponds to $Q_{10}$ values between 1 and 3.3 at 10$^\circ$C, which is the range of values commonly found in the literature (Raich and Schlesinger, 1992). Finally, the $PAR_{1/2}$
parameter is allowed to range over all of the observed values of PAR, which is between 0 and 2000 $\mu$mol m$^{-2}$ s$^{-1}$.

3.3 Results and discussion

3.3.1 Climate

We hypothesize that the primary climate driving variables over seasonal to annual timescales of WLEF NEE variability are SWC, which is strongly related to precipitation and temperature. Although PAR is a primary driving variable over shorter timescales, seasonal and annual sums of this variable generally do not deviate by more than 3-5% from average values. Considerable variability in climate at the WLEF site was observed during the period 1997-2004. Minocqua, which is a National Climatic Data Center reporting station about 30km from WLEF, reported a 30-year annual mean of 4.2$^\circ$C and annual precipitation of 928mm (1975-2004). On average, more than two-thirds of annual precipitation fell in the six-month period from May-October.

Several notable monthly and seasonal temperature anomalies occurred in the period of study. Large temperature anomalies were observed in the December-March period. The most striking temperature anomaly in the April-November period occurred in May of 1997, when the mean monthly temperature was 4.5$^\circ$C below the 30-year average. This contrasts sharply with May of 1998 when the monthly temperature was 3$^\circ$C above the 30-year average. April of 1998 and April of 1999 were also over 2.5$^\circ$C above average. Although June-August of 1998 was near normal in temperature, the
period from September through November of 1998 again averaged about 3°C above normal. During the June-August period, most months were within 2°C of average. The exceptions are July of 1999, which was 2.4°C warmer than normal, and August of 1997 and 2004, which were slightly more than 2°C below normal.

Many monthly precipitation anomalies were observed in Minocqua, especially during the May-October period. The greatest precipitation anomaly was a prolonged drought that occurred in the summer of 1998. For July-September, the three-month precipitation total was 100 mm, or less than 30% of normal. May and July of 1999 were abnormally wet (200% and 150% of normal, respectively). June and July of 2000 were also wetter than average with about 150% of normal rainfall. In 2001, drought conditions were observed, but not as severe as in 1998 with July-September precipitation at about 65% of normal. The 1997-2004 period as a whole was abnormally dry in the months of August and September with the exception of 2002, which is not included in the analysis because fluxes were not available. In August, only 1997 was near normal with the six subsequent years ranging between 20 mm and 80 mm below the normal value of about 120 mm. In September, precipitation every year was below the 30-year average, ranging between 25 mm and 80 mm below the normal of 110 mm.

Temperature and precipitation patterns observed at WLEF were similar to those reported at Minocqua. Notable differences in precipitation occurred in the summer months, when small-scale convective elements comprised a large fraction of precipitation. In July of 1999, observed precipitation at WLEF was even higher than that at Minocqua, totaling over 200 mm. During July-September of 1998, precipitation measured at WLEF totaled about 140 mm, which was higher than at Minocqua but
considerably less than that in the same period in the other four years. In the winter months, the precipitation data from WLEF are not trustworthy because the sensor typically underestimated snowfall (Davis et al., 2003). Patterns of SWC generally reflected the patterns of precipitation observed at Minocqua. SWC tended to decline late in the growing season as a result of transpiration by forest vegetation. However, 1997 is an exception as August values of SWC remained nearly as high as June and July values. Although this appears as an anomaly in the 5-year mean, it may be more representative of the long-term mean SWC because August precipitation values are closest to the 30-year normal in 1997.

3.3.2 Seasonal and interannual variability of gap-filled NEE

Seasonal and annual sums of NEE display a range of variability larger than the estimated random uncertainty (Table 3-1). Seasons are defined by the following months: Winter is January-March and December, spring is April and May, summer is June-August and autumn is September-November. The months defined as winter are generally when soil is frozen or snow-covered; there is little or no GEP during this period and small values of $R_E$. On average, April and May feature steadily increasing GEP and $R_E$, while June-August values are relatively steady. Signs of senescence begin to appear in September, followed by steadily decreasing GEP and $R_E$ through November. Note that the value of the annual NEE sum for 1997 is slightly higher than the published value in Davis et al. (2003) because of minor differences in the preferred NEE and gap-filling algorithms.
Based on the data and error analysis, several important conclusions can be made about the variability of seasonal and annual fluxes between 1997 and 2004. Given a random error of 20 gC m\(^{-2}\) s\(^{-1}\) for two annual sums, a difference of 56 gC m\(^{-2}\) s\(^{-1}\) is

Table 3-1: Sums of gap-filled NEE (units of gC m\(^{-2}\)) for each year and for all four seasons of each year. Uncertainty estimates (1\(\sigma\)) are shown for the annual sum and represent the combined estimate of turbulent sampling and gap-filling uncertainties. The correction for level selection bias has already been applied to the annual NEE sum. Separate daytime and nighttime NEE sums for the full year and all four seasons are also shown.

<table>
<thead>
<tr>
<th>Year</th>
<th>Annual</th>
<th>(\sigma)</th>
<th>DJFM</th>
<th>AM</th>
<th>JJA</th>
<th>SON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NEE (gC m(^{-2}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>47</td>
<td>15</td>
<td>42</td>
<td>18</td>
<td>-94</td>
<td>80</td>
</tr>
<tr>
<td>1998</td>
<td>125</td>
<td>18</td>
<td>45</td>
<td>-21</td>
<td>4</td>
<td>98</td>
</tr>
<tr>
<td>1999</td>
<td>128</td>
<td>20</td>
<td>38</td>
<td>11</td>
<td>24</td>
<td>54</td>
</tr>
<tr>
<td>2000</td>
<td>108</td>
<td>16</td>
<td>47</td>
<td>17</td>
<td>-21</td>
<td>65</td>
</tr>
<tr>
<td>2001</td>
<td>186</td>
<td>17</td>
<td>46</td>
<td>12</td>
<td>69</td>
<td>59</td>
</tr>
<tr>
<td>2003</td>
<td>145</td>
<td>20</td>
<td>59</td>
<td>-4</td>
<td>-7</td>
<td>97</td>
</tr>
<tr>
<td>2004</td>
<td>122</td>
<td>19</td>
<td>29</td>
<td>14</td>
<td>-12</td>
<td>91</td>
</tr>
</tbody>
</table>

| Nighttime NEE (gC m\(^{-2}\)) |      |      |      |     |     |
| 1997 | 441    | 7    | 28   | 31  | 232 | 150 |
| 1998 | 439    | 8    | 28   | 63  | 217 | 131 |
| 1999 | 443    | 11   | 20   | 49  | 266 | 108 |
| 2000 | 436    | 9    | 25   | 59  | 238 | 114 |
| 2001 | 483    | 8    | 27   | 63  | 264 | 130 |
| 2003 | 412    | 10   | 34   | 41  | 218 | 119 |
| 2004 | 410    | 9    | 16   | 47  | 221 | 125 |

| Daytime NEE (gC m\(^{-2}\)) |      |      |      |     |     |
| 1997 | -394   | 8    | 14   | -13 | -326| -69 |
| 1998 | -314   | 10   | 18   | -84 | -214| -33 |
| 1999 | -315   | 12   | 18   | -37 | -242| -54 |
| 2000 | -328   | 8    | 23   | -43 | -259| -49 |
| 2001 | -297   | 10   | 19   | -51 | -194| -71 |
| 2004 | -288   | 9    | 13   | -33 | -234| -34 |
statistically significant between the two sums at the 95% level ($P = 0.05$). Although the differences among many pairs of years are less than $56 \text{ m}^2 \text{ yr}^{-1}$, the range of variability of $140 \text{ gC m}^{-2}$ over seven years is significant ($P < 0.0001$). This contrasts with the results of a study in a nearby forest in lower Michigan, where the range of error exceeds interannual variability (Schmid et al., 2003); however, this study also considered the effects of systematic error such as $u^*$ screening.

On a seasonal scale, there are a number of significant differences in NEE sums among years. Although there is no significant variability during the winter months, the 1998 spring cumulative NEE sum is significantly lower than any other year except for 2003 ($P < 0.05$). Decomposing this sum into nighttime and daytime sums implies that increased GEP is mainly responsible because the nighttime differences (representing $R_E$) are smaller relative to uncertainty. However, the difference between 1997 and 1998 spring values is significant during both nighttime ($P = 0.0005$) and daytime ($P < 1e-6$), implying that $R_E$ is higher in spring 1998 than in 1997. In the summer, 1999 and 2001 nighttime values of NEE are significantly elevated as compared to the other three years ($P < 0.05$). NEE in June-August 1997 is significantly lower than that of any other year ($P = 0.0001$), while 2001 is higher as the only growing season with net release of carbon to the atmosphere. These patterns of significance imply that growing season NEE variability is primarily driven by daytime fluxes. Indeed, the range of daytime NEE ($140 \text{ gC m}^{-2}$) is about twice as great as that of nighttime NEE ($73 \text{ gC m}^{-2}$). 1997 nighttime cumulative sums are similar to 1998 and 2000, while 2001 is similar to 1999. Autumn differences are generally not large, although both daytime and nighttime fluxes in 1997 are significantly larger in magnitude than in several of the other years ($P < 0.05$).
On an annual scale, there is significantly less release during 1997 than in any of the other years (P = 0.03). There is significantly more release in 2001 than in any of the other years except 2003 (P = 0.04). There is also significantly more uptake in 1998 than in 2000. Differences in annual nighttime NEE are not significant except in 2001, in which respiration is elevated compared to the other four years. Much more variability exists in the daytime cumulative NEE. The range of variability in daytime fluxes is over three times that of nighttime fluxes, similar to the pattern of variability in the growing season.

3.3.3 Driving factors of interannual variability

Correlations between seasonal or annual sums of NEE and the primary climate driving variables $T_s$ and SWC are significant ($P < 0.05$) in several cases (Table 3-2). Spring and summer nighttime NEE display significant positive correlations with $T_s$ (P = 0.03). Winter nighttime NEE is negatively correlated with $T_s$ but is not statistically significant (P = 0.12); autumn nighttime NEE displays a weak positive correlation with $T_s$ (P = 0.28). Though not statistically significant, daytime NEE is negatively correlated with $T_s$ in the winter (P = 0.11), spring (P = 0.08), and autumn (P = 0.10) while positively correlated in the summer (P = 0.25). This likely reflects enhanced growth caused during periods of warmer than average temperatures in the transitional seasons, but enhanced respiration in the summer months. For example, positive anomalies in $T_s$ are associated with early leaf-out and enhanced uptake in spring 1998. This effect of springtime temperatures has been shown to be an important driver of interannual variability in boreal
forests (Black et al., 2000; Chen et al., 1999). However, higher $T_s$ in the summer and autumn are associated with decreased uptake. Annual sums of nighttime NEE are strongly correlated with annual mean values of $T_s$ ($P = 0.01$). There is a weaker negative correlation between $T_s$ and annual daytime NEE sums ($P = 0.17$), and virtually no correlation between $T_s$ and total annual NEE sums.

Table 3-2: Correlations between anomalies of NEE sums and (a) soil temperature or (b) soil moisture. Night is defined as all hours when PAR = 0, and day is defined as all hours when PAR > 0. Highlighted values are statistically significant ($P < 0.05$).

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
<th>Whole year</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Night</td>
<td>-0.50</td>
<td>0.73</td>
<td>0.73</td>
<td>0.27</td>
<td>0.81</td>
</tr>
<tr>
<td>Day</td>
<td>-0.54</td>
<td>-0.60</td>
<td>0.31</td>
<td>-0.55</td>
<td>-0.43</td>
</tr>
<tr>
<td>All</td>
<td>-0.67</td>
<td>-0.30</td>
<td>0.57</td>
<td>-0.36</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
<th>Whole year</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Night</td>
<td>0.16</td>
<td>-0.83</td>
<td>0.40</td>
<td>0.75</td>
<td>0.27</td>
</tr>
<tr>
<td>Day</td>
<td>-0.42</td>
<td>0.61</td>
<td>-0.74</td>
<td>-0.57</td>
<td>-0.86</td>
</tr>
<tr>
<td>All</td>
<td>-0.08</td>
<td>0.23</td>
<td>-0.47</td>
<td>-0.03</td>
<td>-0.67</td>
</tr>
</tbody>
</table>

Nighttime NEE is negatively correlated with SWC in the spring ($P = 0.01$), implying that increased moisture limits $R_E$ (Table 3-2b). This may be the result of anaerobic respiration occurring over an increased area of the WLEF footprint, a large fraction of which contains wetlands. A weaker positive correlation between SWC and NEE is observed in the daytime in spring ($P = 0.07$). In the summer, there is a significant negative correlation between SWC and daytime NEE ($P = 0.03$), implying more uptake during moist conditions. This negative correlation persists but becomes weaker during the autumn months ($P = 0.09$). In the nighttime in summer, NEE and SWC are weakly correlated ($P = 0.19$), but a significant positive correlation exists in the nighttime in
autumn ($P = 0.03$). This implies that there is both GEP and $R_E$ enhancement as a result of increased moisture in autumn. Nonparametric models also imply reductions in $R_E$ under conditions of moisture stress (Yi et al., 2004). We hypothesize that wetland areas impose a moisture limitation to respiration during excessively wet conditions (especially during the spring), while upland areas impose a limitation during dry conditions; an “optimal” value of SWC exists somewhere in between. At an annual timescale, both daytime and total NEE display a significant negative correlation with SWC, implying that moisture may be a strong control on interannual variability. However, because SWC and temperature are confounded variables, obtaining a true mechanistic understanding of the effect of climate on respiration requires that we examine the effects of these variables more carefully (Davidson et al., 1998).

### 3.3.4 Modeling results

Posterior parameter distributions are much more tightly constrained than the priors and are non-Gaussian in some cases (Fig. 3-5). Maximum likelihood values and confidence intervals of parameters for the full data assimilation are shown in Table 3-3. Important differences in parameter PDFs occur depending on whether daytime, nighttime, or all data are used. The PDFs of the three temperature-related respiration parameters imply that nighttime respiration is roughly a factor of two larger in the nighttime than during the daytime under typical growing season conditions, whereas the results of the full data assimilation produce estimates of respiration that are roughly an average of daytime and nighttime estimates (Fig. 3-6). This contrasts with the results of
Falge et al. (2002), who report that there is little difference between nighttime respiration and daytime respiration as estimated from light-response relationships at most eddy covariance sites.

Figure 3-5: Probability density functions (PDFs) of interannual model parameters estimated by the MCMC data assimilation technique. The results of three assimilation experiments are shown: the base-case assimilation with all data (thick solid), with nighttime data only (dashed), and with daytime data only (gray).
The maximum likelihood value of $R_{10froz}$ produces reasonable but slightly high-biased estimates of nighttime winter NEE. If only nighttime data are used, $R_{10froz}$ is reduced by a factor of 2, causing better agreement with nighttime winter NEE. If only daytime data are used, the estimate of $R_{10froz}$ is about twice that of the full assimilation. This does not agree with daytime NEE observations in the winter unless there is a GEP contribution during this time that is of the same order of magnitude as respiration, which is not supported by wintertime light-response relationships. Including wintertime data in the assimilation would likely help to better constrain this parameter.

The $R_{10fac}$ and $E_a$ parameters represent the response of respiration to changes in $T_s$ and $T_a$, respectively. The maximum likelihood estimate of this parameter roughly corresponds to a 1 $\mu$mol m$^{-2}$ s$^{-1}$ increase in base respiration for a 4°C in $T_s$. Much smaller values are obtained when only daytime data are used, and larger values occur when

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Best fit</th>
<th>2.5%</th>
<th>97.5%</th>
<th>Prior min</th>
<th>Prior max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{10froz}$</td>
<td>$\mu$mol m$^{-2}$ s$^{-1}$</td>
<td>0.908</td>
<td>0.652</td>
<td>1.216</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>$R_{10fac}$</td>
<td>$\mu$mol m$^{-2}$ s$^{-1}$ K$^{-1}$</td>
<td>0.240</td>
<td>0.207</td>
<td>0.320</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$SWC_{optR}$</td>
<td>vol. frac.</td>
<td>0.296</td>
<td>0.260</td>
<td>0.443</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$a_R$</td>
<td>unitless</td>
<td>0.394</td>
<td>0.082</td>
<td>0.590</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$E_a$</td>
<td>Joules mol$^{-1}$</td>
<td>2.07x10$^4$</td>
<td>1.69x10$^4$</td>
<td>2.38x10$^5$</td>
<td>0</td>
<td>8x10$^4$</td>
</tr>
<tr>
<td>$GEP_{fac}$</td>
<td>$\mu$mol m$^{-2}$ s$^{-1}$ K$^{-1}$</td>
<td>1.31</td>
<td>1.26</td>
<td>1.71</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>$SWC_{optGEP}$</td>
<td>Vol frac</td>
<td>0.304</td>
<td>0.290</td>
<td>0.445</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$a_{GEP}$</td>
<td>unitless</td>
<td>0.568</td>
<td>0.412</td>
<td>0.719</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$PAR_{1/2}$</td>
<td>$\mu$mol m$^{-2}$ s$^{-1}$</td>
<td>458</td>
<td>408</td>
<td>534</td>
<td>0</td>
<td>2000</td>
</tr>
</tbody>
</table>

Table 3-3: Parameters used in the interannual NEE model. The best fit solutions are the value of the parameter at which the likelihood function is maximized. The 2.5 and 97.5 percentiles of the posterior parameter PDFs represent the bounds of the 95% posterior confidence intervals derived using Markov Chain Monte Carlo. In all cases, priors were uniformly distributed between the prior min and max.
nighttime data are used. The temperature sensitivity $E_a$ is somewhat low, corresponding to a $Q_{10}$ of about 1.4 at 10°C. When nighttime data are used, the corresponding $Q_{10}$ is about 1.8; when daytime data are used, the corresponding $Q_{10}$ is also slightly higher at 1.5.

The parameters representing the soil moisture dependence of respiration were not as well constrained. The 95% confidence interval of $SWC_{optR}$ covers the entire upper half of the prior range of observed SWC at WLEF when either all data or only nighttime data are used. When daytime data are used, $SWC_{optR}$ is more sharply constrained around a lower maximum likelihood value of about 0.18. Because the landscape around WLEF is a mixture of wetland and upland ecosystems, the breadth of the $SWC_{optR}$ parameter may reflect the mixed landscape and may indicate the shortcomings of using a single SWC point measurement. The broad 95% confidence intervals of the $a_R$ parameter in all three assimilations demonstrate that the soil moisture dependence of respiration is poorly constrained.

The temperature dependence parameter $GEP_{fic}$ is tightly constrained. The maximum likelihood solution of this parameter is slightly lower when only daytime data are used compared to when all data are used. This reflects the fact that daytime respiration estimates are lower, meaning that a smaller estimate of GEP is required to match the observed NEE (Fig. 3-6). The half saturation constant $PAR_{1/2}$ is also higher when only daytime data are used, having a similar effect of reducing GEP estimates.
The soil moisture parameters $a_{\text{GEP}}$ and $\text{SWC}_{\text{optGEP}}$ behave similarly as the equivalent respiration parameters. Because the edge of the $\text{SWC}_{\text{optGEP}}$ PDF hits the prior maximum constraint, it is unclear whether there is an optimal value of SWC or if GEP continues to increase as a function of SWC. Possible mechanisms for the reduction of GEP at high values of SWC are not understood, and the model formulation was chosen only to enable direct comparison between respiration and GEP parameters. The maximum likelihood values of $\text{SWC}_{\text{optGEP}}$ and $a_{\text{GEP}}$ do indicate a reduction of GEP at high values of SWC, which may be an artifact of structural uncertainty in the model; if $f_{\text{GEP}}(T_s)$ is erroneous in predictions of spring phenology, the fact that the highest values of SWC generally occur prior to leafout may play a role in this effect. An underestimation of $R_E$ at high values of SWC may also be the cause of this effect. Sharp reduction of GEP is also predicted during extremely dry conditions. We note that this effect may be overestimated if $f_{\text{GEP}}(T_s)$ is erroneous in predictions of senescence; the lowest values of

Figure 3-6: Dependence of $R_{10}$ on soil temperature as estimated by the interannual model (a) for the base-case assimilation (thick solid), nighttime-only assimilation, and (dashed) daytime-only assimilation (gray). Dependence of GEP at a value of $\text{PAR} = 1000 \, \mu\text{mol m}^{-2} \, \text{s}^{-1}$. 
SWC observed at WLEF generally occur around the time of senescence. Further research regarding the moisture dependence of GEP at WLEF is warranted, especially given the high correlation between SWC and daytime NEE on an annual timescale.

The interannual model reproduces synoptic cycles and the mean seasonal cycle of NEE at WLEF reasonably well when all data are assimilated (Fig. 3-7). Synoptic NEE variability is slightly underestimated, implying that there are driving factors of this variability not captured by the model. This may be related to variability in the size and orientation of the flux footprint combined with heterogeneity in the landscape. The model systematically overestimates wintertime respiration, which is a result of omitting wintertime data from the data assimilation and of neglecting the effects of snow cover. Uptake is underestimated in the springtime but overestimated late in the growing season and in the autumn. This is likely a result of assuming that the response of NEE to $T_z$ is identical during these seasons. To improve accuracy, leafout and senescence should be modeled as separate processes.

Reproducing the observed patterns of interannual variability in annual sums of NEE is more difficult, and the model is only moderately successful (Fig. 3-8). The model reproduces interannual variability the two extreme years of 1997 (high uptake) and 2001 (high release), but underestimates the range of variability by about a factor of two. Interannual variability among the other five years is poorly modeled, but the observed range of variability is not statistically significant ($P = 0.30$). The model is also moderately successful in reproducing interannual variability in daytime and nighttime NEE sums. The magnitudes of both daytime and nighttime sums are systematically underestimated by the model, reflecting biases in respiration parameters (Fig. 2-6).
Because the interannual model is based on standard gap-filling equations, this analysis indicates that using these gap-filling equations may produce biased estimates of GEP and \( R_E \). This bias occurs when extrapolating nighttime NEE relationships to estimate daytime respiration as a function of temperature.

On seasonal timescales, the interannual model was able to reproduce the low values of observed nighttime NEE in spring of 1997 and high values in the spring of 1998, the key source of variability in this period (Table 3-4). Variability in daytime and total spring NEE was poorly modeled. The correlation between observed and modeled summer nighttime NEE variability is significant (\( P = 0.03 \)), and the elevated values of observed \( R_E \) in the growing season of 1999 are accurately modeled. Daytime summer NEE variability is not as well modeled. Although there is a correlation between
modeled and observed total summer NEE (P = 0.06), the model underestimates the range of variability by about a factor of four. Variability in autumn NEE was poorly modeled. Similar MCMC analyses also conclude that interannual NEE variability is difficult to capture in a model. Braswell et al. (2005) used MCMC to calibrate a somewhat more complex model to eddy covariance data from Harvard Forest, and noted that this model did not perform as well on an interannual timescale as at diurnal and seasonal timescales. Additional years of data would further constrain the modeled NEE at WLEF especially during abnormally moist or dry conditions. For example, the data record at WLEF is

<table>
<thead>
<tr>
<th>Season</th>
<th>Time of day</th>
<th>Slope (gC m⁻²)</th>
<th>Obs.mean (gC m⁻²)</th>
<th>Range (gC m⁻²)</th>
<th>Model bias (gC m⁻²)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr-Nov</td>
<td>All</td>
<td>0.43</td>
<td>79</td>
<td>135</td>
<td>-3</td>
<td>0.63</td>
</tr>
<tr>
<td>Apr-Nov</td>
<td>Night</td>
<td>0.96</td>
<td>412</td>
<td>79</td>
<td>-37</td>
<td>0.72</td>
</tr>
<tr>
<td>Apr-Nov</td>
<td>Day</td>
<td>0.63</td>
<td>-333</td>
<td>116</td>
<td>34</td>
<td>0.57</td>
</tr>
<tr>
<td>Spring (AM)</td>
<td>All</td>
<td>-0.10</td>
<td>7</td>
<td>40</td>
<td>8</td>
<td>0.02</td>
</tr>
<tr>
<td>Spring (AM)</td>
<td>Night</td>
<td>0.81</td>
<td>51</td>
<td>32</td>
<td>4</td>
<td>0.61</td>
</tr>
<tr>
<td>Spring (AM)</td>
<td>Day</td>
<td>0.33</td>
<td>-44</td>
<td>71</td>
<td>4</td>
<td>0.18</td>
</tr>
<tr>
<td>Summer (JJA)</td>
<td>All</td>
<td>0.25</td>
<td>-5</td>
<td>163</td>
<td>-9</td>
<td>0.63</td>
</tr>
<tr>
<td>Summer (JJA)</td>
<td>Night</td>
<td>0.65</td>
<td>237</td>
<td>49</td>
<td>-49</td>
<td>0.86</td>
</tr>
<tr>
<td>Summer (JJA)</td>
<td>Day</td>
<td>0.31</td>
<td>-242</td>
<td>132</td>
<td>39</td>
<td>0.46</td>
</tr>
<tr>
<td>Autumn (SON)</td>
<td>All</td>
<td>0.14</td>
<td>778</td>
<td>44</td>
<td>-2</td>
<td>0.08</td>
</tr>
<tr>
<td>Autumn (SON)</td>
<td>Night</td>
<td>0.35</td>
<td>125</td>
<td>42</td>
<td>8</td>
<td>0.30</td>
</tr>
<tr>
<td>Autumn (SON)</td>
<td>Day</td>
<td>0.29</td>
<td>-47</td>
<td>49</td>
<td>-10</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Table 3-4: Goodness of fit statistics for the interannual model as a function of time of day and time of year. Slopes are determined by performing a linear fit of the model as a function of observed NEE; slopes are less than one in all cases, indicating that the model underestimates interannual variability. Observed means and ranges are determined using gap-filled NEE from the period 1997-2004 excluding 2002. Model bias indicates the deviation of the model mean from the observed mean for the given time of day and year. Highlighted values of R² are statistically significant (P < 0.05).
Figure 3-8: Modeled vs. observed annual NEE (a), annual nighttime NEE (b), and annual daytime NEE (c). The interannual model systematically underestimates the magnitudes of daytime and nighttime fluxes. Vertical error bars are derived from the MCMC parameter uncertainty estimates, and horizontal error bars represent observational uncertainty estimates. Error bars and fit statistics are shown in table 4.
dominated by dry growing seasons with respect to the 30-year mean; wetter growing seasons may help to better constrain the moisture dependence of NEE and to help discern the effects of moisture from the effects of phenology.

More research is also necessary to refine the model in order to reduce the observed structural uncertainties. One structural shortcoming of the respiration submodel is that it fails to capture the differences between nighttime and daytime respiration. The primary advantages of the model used in this study is that: 1) it uses a single set of parameters to predict hourly fluxes, which can be aggregated to any timescale, and 2) derived parametric uncertainties can be used to make probabilistic predictions of fluxes. Such models could eventually be used for gap-filling of missing data, with gap-filling uncertainties derived from model parametric uncertainties.

3.3.5 Conclusion

We analyzed seven years of WLEF NEE observations from the period 1997-2004. Systematic errors in NEE sums were corrected when possible, and random errors related to turbulent variability and uncertainty in gap-filling were estimated. Random errors were compared to the range of interannual variability of NEE to test for statistical significance of this variability. Seasonal and annual NEE sums were correlated with two climate variables: soil moisture and temperature. The gap-filling model was extended to produce an interannual model with the goal of retrieving a single set of parameters in order to describe NEE variability across all timescales from diurnal to interannual at WLEF. We used the Markov Chain Monte Carlo technique in order to
assimilate hourly observations of NEE into this model and to estimate probability density functions of all model parameters. Parameter PDFs were then used to estimate NEE and uncertainties, which we compared to observations. The conclusions of this analysis are the following:

1. The range of interannual variability in NEE, RE and GEP at WLEF exceeds the random error due to turbulent variability and gap-filling and is statistically significant at the 95% level. Although systematic errors are potentially much larger than the estimated random error, we attempt to correct for these errors when possible and assume that remaining errors do not affect estimates of NEE variability. The choice of u* threshold does not affect patterns of interannual variability except at very high or low values. WLEF is a source of carbon to the atmosphere from 1997-2004 regardless of the choice of u* threshold.

2. Statistically significant correlations are observed between NEE variability and anomalies of temperature and soil moisture at seasonal and annual timescales in many cases. At an annual timescale, NEE correlates with soil moisture but not temperature. A better model formulation of moisture dependence, perhaps including water table depths or soil moisture from multiple sites in the WLEF footprint, may result in a more accurate representation of interannual variability.

3. The range of interannual variability is greater in the spring and summer than in autumn and winter, and greater during the day than at night. Spring variability is related primarily to smoothed soil temperature, which is correlated to the date of leafout. Summer variability may be caused by temperature and moisture. The
competing effects of temperature and moisture tend to limit autumn variability, and wintertime variability is small despite large variations in temperature.

4. Estimated interannual model parameters are generally not constrained by prior estimates. Parameter PDFs are not necessarily Gaussian, and these PDFs change depending on whether daytime, nighttime or all NEE data are used in the assimilation. This reflects significant differences between estimated nighttime and daytime respiration for the same given soil moisture and temperature.

5. The interannual model reproduces interannual variability in annual and growing season NEE with moderate success. The model is unsuccessful in reproducing interannual variability in the transition seasons of spring and autumn. In all cases, the model underestimates the range of interannual variability. Further research is necessary to better understand the driving factors of interannual variability at WLEF.
4.1 Introduction

Atmospheric CO₂ concentrations are currently increasing at an average rate of roughly 1.5 parts per million (ppm) per year, which is about one half of the rate of fossil fuel emissions. The other half of fossil fuel emissions is absorbed by a surface CO₂ sink, which consists of the oceans and terrestrial ecosystems. This sink of CO₂ exhibits significant interannual variability, with a 5 GtC range in the 1990s. The magnitude of the total sink is well constrained, as it is determined by the difference between two well-known quantities: annual global CO₂ increase and fossil fuel emissions. Although inversion models indicate that the terrestrial sink is smaller than or equal in magnitude to the oceanic sink (Gurney et al., 2002; Rodenbeck et al., 2003b), as much as two thirds of total sink variability has been attributed to the terrestrial biosphere because the strength of the oceanic sink is likely more consistent on an interannual timescale (Bousquet et al., 2000; Rodenbeck et al., 2003a).

Net carbon uptake by the terrestrial carbon cycle is driven by two large ecosystem fluxes: net primary productivity (NPP) and heterotrophic respiration (R_H). Small relative

---

3 This chapter is an early draft of an article by Ricciuto, D.M., McInerney, D., Davis, K.J. and Keller, K., Improving model estimates of spatial and temporal variations in CO₂ fluxes: A multi-site data assimilation of TRIFFID.
variations in these gross fluxes can cause large variations in net uptake. Climate, mainly through temperature, radiation and precipitation, has been shown to be a primary driver of interannual variability in both NPP (Nemani et al., 2002) and respiration (Raich et al., 2002). Terrestrial CO₂ fluxes over large regions display statistically significant correlations with climate indices such as the Southern Oscillation and Arctic Oscillation (Potter et al., 2003). Because of the sensitivity of these fluxes to climate, long-term climate change is likely to have a significant impact on the magnitude of the terrestrial carbon sink in the future.

A number of competing carbon-climate feedback mechanisms cause large uncertainty in predictions of future CO₂ fluxes. For example, two recent studies used different coupled GCM-carbon cycle models to make predictions of CO₂ concentrations and climate for the next 100 years. One study (Cox et al., 2000) predicts a strong positive feedback between respiration and temperature, causing the terrestrial biosphere to switch from a sink to a source of CO₂ after 2050. A second study predicts that the terrestrial biosphere remains a carbon sink through the year 2100 (Dufresne et al., 2002). Recently, ten such coupled carbon-climate models were compared (Friedlingstein et al., 2006), with a large structural uncertainty in the predicted fates of the terrestrial CO₂ sink: by 2100, there is a 16 GtC yr⁻¹ range in predicted terrestrial CO₂ fluxes. These models also differ strongly in their predictions of interannual variability of terrestrial fluxes over the past 50 years. This uncertain response of the terrestrial carbon cycle to future climatic change remains an important problem in predicting future climate. Alleviating this problem requires developing an improved mechanistic understanding that can lead to improved models of the terrestrial carbon cycle.
One way to improve models is to constrain them with environmental observations of CO$_2$ and climate. Because fully coupled GCM/carbon cycle models are computationally expensive to evaluate and data assimilation techniques demand large numbers of model runs, the focus of this effort to date has been so far on relatively simple models. Knorr and Heimann (1995) developed the Simple Diagnostic Biosphere Model (SDBM), in which both NPP and $R_H$ are modeled from input variables of Normalized Difference Vegetation Index (NDVI), NPP, temperature, precipitation and solar radiation. Two parameters were optimized: $\varepsilon$ (light use efficiency) and $Q_{10}$ (factor of respiration increase for a $10^\circ C$ rise in temperature). This model was found to simulate the annual cycle in atmospheric CO$_2$ with a skill several times greater than more complex ecosystem models with many parameters that were not tuned to observations (Heimann et al., 1998). This model was not used to predict interannual variability and is of limited use in predicting future fluxes because it requires satellite-derived NDVI as input. Another modeling study involved fitting sixteen ecosystem parameters to produce a reasonably accurate prediction of variations in the strength of the global terrestrial carbon sink (Vukicevic et al., 2001). This study suggests that lagged temperature effects and nutrient cycling are necessary to include in a model of interannual variability, but did not consider other climate variables such as precipitation or allow for oceanic flux variability. A recent study by Rayner et al. (2005) involved an assimilation of 57 parameters in the Biosphere Energy Transfer Hydrology Model (BETHY) in conjunction with the TM2 transport model using CO$_2$ concentration observations from GLOBALVIEW-CO$_2$ as constraints. The optimized model reproduces the observed global CO$_2$ growth rate reasonably well, but the method used in this study is prone to misconvergence.
None of these data assimilation studies incorporate eddy covariance observations. Eddy covariance towers measure continuous fluxes of CO₂ at half-hourly or hourly timesteps in addition to a number of meteorological variables such as air and soil temperature, precipitation and radiation. Currently, there are over 200 such towers spread throughout the world (Baldocchi et al., 2001). Eddy covariance measurements have been used to optimize terrestrial models, improving predictions of carbon fluxes at Harvard Forest (Braswell et al., 2005), Niwot Ridge (Sacks et al., 2006), a site in Kansas (Knorr and Kattge, 2005), and WLEF in northern Wisconsin (Ricciuto et al., 2006). Although model predictions improve from diurnal to seasonal timescales, interannual variability is usually poorly modeled even after parameters are optimized to improve the model fit.

In this study, we optimize parameters in the TRIFFID (Top-down Representation of Interactive Foliage and Flora Including Dynamics) model for five eddy covariance sites in eastern U.S. forests with long data records. These sites contain a combined 37 site-years of data. TRIFFID is chosen because it is a fully prognostic dynamic vegetation model that has been coupled with multiple GCMs (Cox et al., 2000; Matthews et al., 2005a) to make predictions of climate-carbon cycle feedbacks. Evaluating the performance of TRIFFID in present-day is of considerable relevance to assessing potential biases of future predictions. For computational feasibility, we limit the optimization to model parameters that are expected to be most sensitive to an assimilation of carbon flux data. We optimize model parameters individually for each site in order to evaluate the ability of the model to reproduce observed fluxes from hourly to interannual timescales. Optimized parameters are compared across sites to assess spatial coherence.
We then perform a joint optimization of all five sites, obtaining a single set of model parameters to predict carbon fluxes across all site-years. Variation among sites is assumed to be caused solely by variations in climate forcing, canopy height, PFT distribution, soil carbon and leaf nitrogen content. We then evaluate the ability of the jointly optimized model to reproduce both interannual and cross-site variability. We test the following key hypotheses:

1. The parameter space is nonconvex (contains multiple local maxima), necessitating a global optimization method. Gradient-based optimization methods are likely to misconverge to local rather than global optima.

2. The single-site parameter optimizations improve correlations of both monthly fluxes (seasonal variability) and annual fluxes (interannual variability) at all five sites.

3. A joint optimization of all sites, in which a single set of parameters is used to predict fluxes, improves the estimate of cross-site variability.

4. Using observed soil moisture rather than modeled soil moisture produces improves correlations of interannual fluxes.
4.2 Materials and Methods

4.2.1 Site descriptions

The five sites chosen for the optimization are WLEF in northern Wisconsin, the University of Michigan Biological Station (UMBS) in northern lower Michigan, Howland Forest in southern Maine, Harvard Forest in central Massachusetts, and Morgan Monroe in Indiana (Figure 4-1). All five sites compute net ecosystem exchange of CO₂ (NEE) using similar instrumentation and data processing techniques (Baldocchi et al., 2001). These sites all have forested footprints with varying mixtures of deciduous and coniferous species. Our implementation of the TRIFFID model requires specification of canopy height and the fractions of broadleaf and needleleaf trees in the tower footprints. These input data, in addition to site locations and references, are summarized in Table 1.

WLEF is located in the Park Falls Ranger District of the Chequamegon National Forest about 15km east of Park Falls, Wisconsin in a region of low relief: typical hilltop to valley elevation change is about 20m over horizontal scales of a few hundred meters. WLEF is unique because of its height and multiple flux measurement levels (30 m, 122 m, and 396 m). A “preferred NEE” product is compiled using different levels depending on availability and micrometeorological conditions (Davis et al., 2003). WLEF is also unique among the five chosen sites because its footprint contains wetlands. Soil moisture depends on elevation, with saturated lowlands (wetland) and unsaturated uplands. The landscape is managed, with scattered areas of thinning and clear-cuts mainly in the upland regions. The whole region was heavily logged around the beginning of the 20th century. Key species of trees found in the upland regions include aspen,
balsam fir, sugar maple, red maple, basswood, red pine, paper birch, yellow birch, and white spruce (Davis et al., 2003). The average stand age is 60-80 years. Wetlands comprise about 40% of the surrounding landscape and include alder, cedar, tamarack, and black spruce stands. Mackay et al. (2002) document the regional forest cover in more detail. Maximum canopy height in the region is about 25 m, but wetlands tend to have substantially lower canopies (~4 meters) as do young aspen stands. Soils are sandy loam

Figure 4-1: Map of the analyzed Ameriflux eddy covariance tower locations. Sites with long records (≥ 5 years) are selected to assess the effectiveness of the model in reproducing observed interannual variability. Details about the site characteristics are given in Table 1.
and are mostly glacial outwash deposits. NEE data from 1997-2004 are used in this study, with the exception of 2002 when data were not available.  

The UMBS measurement site lies near the northern extent of Michigan’s lower peninsula at an elevation of 234 m. Flux measurements are made at both 34 and 46 m, but 46 m fluxes are used in this study. The primary species composition of forests within the tower footprint are bigtooth aspen, trembling aspen, red and sugar maples, red oak, birch and beech (Schmid et al., 2003). There is also a small component of white pine.

<table>
<thead>
<tr>
<th>Site</th>
<th>Latitude (°N)</th>
<th>Longitude (°W)</th>
<th>Canopy ht (m)</th>
<th>BL (%)</th>
<th>NL (%)</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvard Forest</td>
<td>42.538</td>
<td>72.172</td>
<td>24.0</td>
<td>80.0</td>
<td>20.0</td>
<td>(Barford et al., 2001)</td>
</tr>
<tr>
<td>Howland Forest</td>
<td>45.204</td>
<td>68.74</td>
<td>19.5</td>
<td>11.0</td>
<td>89.0</td>
<td>(Hollinger et al., 2004; Hollinger et al., 1999)</td>
</tr>
<tr>
<td>WLEF</td>
<td>45.946</td>
<td>90.272</td>
<td>15.0</td>
<td>70.0</td>
<td>30.0</td>
<td>(Davis et al., 2003; Ricciuto et al., 2006)</td>
</tr>
<tr>
<td>UMBS</td>
<td>45.560</td>
<td>84.714</td>
<td>20.0</td>
<td>98.0</td>
<td>2.0</td>
<td>(Schmid et al., 2003)</td>
</tr>
<tr>
<td>Morgan Monroe</td>
<td>39.323</td>
<td>-86.413</td>
<td>26.0</td>
<td>85.0</td>
<td>15.0</td>
<td>(Schmid et al., 2000)</td>
</tr>
</tbody>
</table>

The stand age is roughly 80-90 years following a clearcut in 1880, and a series of fires ending in 1923. The canopy height averages about 22 m. The mean annual temperature is about 6.2°C with annual precipitation of about 750 mm. Data from 1999-2003 are used in the optimization.
Howland forest is located about 35 miles north of Bangor Maine at an elevation of 60 meters. Fluxes are measured at a height of 29 m over a canopy height of roughly 20 m. Forest stands are dominated by red spruce and eastern hemlock, and also include balsam fir, white pine, northern white cedar, and about 11% hardwoods (red maple, and paper birch). The terrain is mostly flat or gently rolling, but there are large variations in soil drainage. Mean annual temperature is about 6.1°C with mean annual precipitation about 988 mm (Hollinger et al., 1999). The region did experience disturbance in the form of a record ice storm in January of 1998, causing an increase in the amount of coarse woody debris and larger respiration rates (Hollinger et al., 2004). We do not attempt to model this effect. Data from 1996-2003 are used in the optimization.

Morgan Monroe State Forest (MMSF) is located in south–central Indiana at an elevation of 275 m. The terrain is dominated by ridge/ravine topography with local elevation differences of 60 m. Soils are generally well-drained, and the tower footprint is covered by twenty-nine primarily deciduous species. The age of the forest is between 60 and 90 years, with an average canopy height is roughly 27 m. Flux measurements are made at 34 m and 46 m. The dominant species are sugar maple, yellow poplar, sassafras, white oak, and black oak (Schmid et al., 2000). The mean annual temperature is the warmest of the five sites at 12.4°C, and the mean precipitation is 1030 mm as measured at the tower between 1999 and 2003.

Harvard forest is located in central Massachusetts, and has the longest site record in the Ameriflux network of 12 years (1992-2003). The site elevation is 303 m with gentle relief of about 30 m. The canopy height of 24 m, and fluxes are measured at 30 m above ground level. The stand age is roughly 60 years. The dominant species in Harvard
forest are red oak, red maple, black birch, white pine, and hemlock. The mean annual temperature as measured at the tower is about 6.5°C with mean annual precipitation roughly 1000 mm. This site displays significant interannual variability (Barford et al., 2001), and has also been the focus of a data assimilation study (Braswell et al., 2005).

4.2.2 TRIFFID model description

The TRIFFID (Top-Down Representation of Interactive Foliage and Flora Including Dynamics) is a dynamic global vegetation model (DGVM) that includes five plant functional types (PFTs): Broadleaf tree, needleleaf tree, C₃ grass, C₄ grass and shrub. Soil carbon, canopy structure and fractional coverage of these five PFTs are updated based on carbon fluxes, which are calculated using the Met Office Surface Exchange Scheme (MOSES). MOSES is a single soil-layer land surface scheme and is described by Cox et al. (1999). TRIFFID was coupled to GCMs to make future climate predictions by Cox et al. (2000) and Jones et al. (2001), and is described in detail by Cox et al. (2001). This version of TRIFFID/MOSES was also coupled with the University of Victoria Earth System Climate Model (Matthews et al., 2005b). In this study, we only consider the broadleaf and needleleaf tree PFTs and therefore do not discuss aspects of the model that deal with C₄ photosynthesis. We add an improved representation of model phenology, discussed in 4.2.2.3.

The flow of carbon in the TRIFFID model is relatively simple, with three vegetation pools per plant functional type and one soil carbon pool (Fig. 4-2). Net primary productivity (NPP) is calculated separately for each PFT. The details of this
calculation follow below. This NPP is allocated to leaf, wood and root pools for each PFT. All three of these pools turn over at prescribed rates, increasing the soil carbon pool from which the rate of heterotrophic respiration is calculated. For the purposes of this assimilation, the values of the soil, leaf, wood and root pools are held constant and vegetation dynamics are not considered. Table 4-2 is a list of model parameters that includes published values and upper and lower bounds used in the optimizations.

Figure 4-2: Schematic diagram of the flow of carbon in the TRIFFID model. Gross primary productivity (GPP), net primary productivity (NPP) and autotrophic respiration ($R_a$) are calculated separately for the needleleaf (NL) and broadleaf (BL) functional types. NPP is allocated to leaf, wood and root pools, which all turn over into a single soil carbon pool. Heterotrophic respiration ($R_H$) is a function of soil carbon, soil temperature and soil moisture. NPP is the balance between GPP and $R_a$, and net ecosystem exchange (NEE) is the balance between NPP and $R_H$. 
Table 4-2: TRIFFID model parameters optimized in this study.

<table>
<thead>
<tr>
<th>Parameter symbol</th>
<th>Name</th>
<th>Units</th>
<th>Default value</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>Soil carbon</td>
<td>kg C m$^{-1}$</td>
<td>10.0</td>
<td>0.0</td>
<td>30.0</td>
</tr>
<tr>
<td>Q$_{10H}$</td>
<td>Heterotrophic Q$_{10}$</td>
<td>dimless</td>
<td>2.0</td>
<td>1.0</td>
<td>4.0</td>
</tr>
<tr>
<td>$\theta_{opt}$</td>
<td>Optimal heterotrophic SWC</td>
<td>Vol frac.</td>
<td>0.2</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$\theta_{fac}$</td>
<td>SWC dependence of R$_H$</td>
<td>Dimless</td>
<td>0.8</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$\theta_c$</td>
<td>Critical SWC</td>
<td>Vol frac.</td>
<td>0.30</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$\theta_w$</td>
<td>Wilting SWC</td>
<td>Vol frac.</td>
<td>0.13</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>n$_{0BL}$</td>
<td>BL Top leaf nitrogen content</td>
<td>kg N (kg C)$^{-1}$</td>
<td>0.036</td>
<td>0.01</td>
<td>0.20</td>
</tr>
<tr>
<td>n$_{0NL}$</td>
<td>NL Top leaf nitrogen content</td>
<td>kg N (kg C)$^{-1}$</td>
<td>0.030</td>
<td>0.01</td>
<td>0.20</td>
</tr>
<tr>
<td>$\alpha_{BL}$</td>
<td>BL Quantum efficiency</td>
<td>mol C (mol PAR)$^{-1}$</td>
<td>0.06</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>$\alpha_{NL}$</td>
<td>NL Quantum efficiency</td>
<td>mol C (mol PAR)$^{-1}$</td>
<td>0.06</td>
<td>0.01</td>
<td>0.25</td>
</tr>
<tr>
<td>Q$_{10VM}$</td>
<td>T-sensitivity of V$_m$</td>
<td>Dimless</td>
<td>2.0</td>
<td>1.0</td>
<td>4.0</td>
</tr>
<tr>
<td>T$_{lowBL}$</td>
<td>Lower T$_{leaf}$ for BL phot.</td>
<td>°C</td>
<td>-5.0</td>
<td>-40.0</td>
<td>10.0</td>
</tr>
<tr>
<td>T$_{lowNL}$</td>
<td>Lower T$_{leaf}$ for NL phot.</td>
<td>°C</td>
<td>-15.0</td>
<td>-40.0</td>
<td>10.0</td>
</tr>
<tr>
<td>T$_{uppBL}$</td>
<td>Upper T$_{leaf}$ for BL phot.</td>
<td>°C</td>
<td>33.0</td>
<td>10.0</td>
<td>50.0</td>
</tr>
<tr>
<td>T$_{uppNL}$</td>
<td>Upper T$_{leaf}$ for NL phot</td>
<td>°C</td>
<td>28.0</td>
<td>10.0</td>
<td>50.0</td>
</tr>
<tr>
<td>D$_c$</td>
<td>Critical humidity deficit</td>
<td>Kg (kg)$^{-1}$</td>
<td>0.09</td>
<td>0.01</td>
<td>0.50</td>
</tr>
<tr>
<td>R$_g$</td>
<td>Growth resp. coefficient</td>
<td>Dimless</td>
<td>0.25</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>R$_{dc}$</td>
<td>Dark respiration coefficient</td>
<td>Dimless</td>
<td>0.015</td>
<td>0.003</td>
<td>0.075</td>
</tr>
<tr>
<td>Q$_{10RD}$</td>
<td>Dark respiration Q$_{10}$</td>
<td>Dimless</td>
<td>2.0</td>
<td>1.0</td>
<td>4.0</td>
</tr>
<tr>
<td>T$_{off}$</td>
<td>Crit. T$_{leaf}$ for leaf-off</td>
<td>°C</td>
<td>0.0</td>
<td>-30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>$\beta_{off}$</td>
<td>Crit. moisture for leaf-off</td>
<td>Dimless</td>
<td>0.90</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>Budburst/leafoff coeff</td>
<td>yr$^{-1}$</td>
<td>20.0</td>
<td>1.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
4.2.2.1 Photosynthesis

In this scheme, photosynthesis follows existing models of leaf-level C$_3$ photosynthesis (Collatz et al., 1991). The rate of gross leaf photosynthesis is calculated as a function of three potentially limiting factors. The first factor, $W_c$, applies when the enzyme activity of RuBisCO is limiting according to:

$$W_c = V_m \frac{c_i - \Gamma}{c_i + K_c (1 + O_a / K_o)}$$ \hspace{1cm} (4.1)

Here, $c_i$ is the internal leaf CO$_2$ partial pressure, and $\Gamma$ is the value of $c_i$ at which photosynthesis balances photorespiration, or the “photorespiration compensation point”. This parameter has a weak temperature dependence. The value of $c_i$ is modeled by assuming the following parameterization:

$$\frac{c_i - \Gamma}{c_c - \Gamma} = F_0 \left(1 - \frac{D_o}{D_c}\right)$$ \hspace{1cm} (4.2)

where $c_c$ is the leaf surface CO$_2$ partial pressure, $D_o$ is the humidity deficit calculated by MOSES, $D_c$ is the critical humidity deficit parameter, and $F_0$ is a parameter relating $c_i$ to $c_c$ when $D_o = D_c$. In this study, $D_c$ is optimized while $F_0$ is set to a constant value of 0.8. In Eq. 4.1, $O_a$ is the partial pressure of atmospheric oxygen, and $K_c$ and $K_o$ are Michaelis-Menton constants for CO$_2$ and O$_2$, and depend on temperature. $V_m$ is the maximum rate of carboxylation of Rubisco, which depends on leaf temperature:

$$V_m = \frac{V_{max} Q_{10/3}^{0.1(T_e - 25)}}{[1 + \exp(0.3(T_e - T_{app}))][1 + \exp(0.3(T_{low} - T_e))]}$$ \hspace{1cm} (4.3)
Here, $V_{\text{max}}$ is a PFT-dependent parameter describing the maximum rate of carboxylation of Rubisco at a temperature of 25°C. $V_{\text{max}}$ is linearly dependent on the top-leaf nitrogen content $n_{lo}$ and is a PFT-specific parameter. $Q_{10VM}$ describes the temperature sensitivity of this parameter, and $T_{\text{upp}}$ and $T_{\text{low}}$ are PFT-specific parameters that act to limit $V_m$ when leaf temperatures are too warm or too cold, respectively.

The second factor, $W_l$ applies when light is the limiting factor for photosynthesis:

$$W_l = \alpha(1 - \omega)I_{\text{par}} \left[ \frac{c_i - \Gamma}{c_i + 2\Gamma} \right]$$  \hspace{1cm} (4.4)

The parameter $\alpha$ is PFT-specific and is the quantum efficiency. The parameter $\omega$ is the leaf scattering coefficient for PAR, and $I_{\text{par}}$ is the incident photosynthetically active radiation that is specified as an input variable. The third factor, $W_c$ applies when photosynthesis is limited by transport of photosynthetic products. For C$_3$ plants, this is:

$$W_c = 0.5V_m$$  \hspace{1cm} (4.5)

The rate of gross photosynthesis $W$ is calculated by taking a smoothed minimum of the limiting rates $W_c$, $W_l$ and $W_e$. This involves two co-limitation coefficients, which are not optimized in this study.

To calculate gross primary productivity (GPP), the rate of dark respiration $R_d$ is subtracted from the quantity $W$, and a moisture dependence is added. $R_d$ is given by the following relationship:

$$R_d = R_{dec}V_{\text{max}}^{0.1(T_c - 25)}$$  \hspace{1cm} (4.6)
At 25°C, \(R_d\) is a fraction of \(V_{\text{max}}\) given by the coefficient \(R_{dc}\). \(Q_{10RD}\) is the temperature sensitivity of dark respiration, which we optimize. GPP is then given by:

\[
GPP = \beta(W - R_d)
\]

The moisture stress factor \(\beta\) is a function of the soil water content \(\theta\):

\[
\beta = \begin{cases} 
1, & \theta > \theta_c \\
\frac{\theta - \theta_w}{\theta_c - \theta_w}, & \theta_w < \theta \leq \theta_c \\
0, & \theta \leq \theta_w
\end{cases}
\]

\(\theta_c\) and \(\theta_w\) are the critical and wilting values of soil moisture, respectively.

### 4.2.2.2 Respiration

Heterotrophic and autotrophic respiration are modeled separately in TRIFFID. Microbial soil respiration depends on soil carbon content, soil moisture and soil temperature:

\[
R_h = \kappa_s C_s f_{\theta} Q_{10RH}^{0.1(T_c - 25^\circ C)}
\]

\[
f_{\theta} = \begin{cases} 
1 - \theta_{\text{fac}} \left(\theta - \theta_{\text{opt}}\right), & \text{for } \theta > \theta_{\text{opt}} \\
(1 - \theta_{\text{fac}}) + \theta_{\text{fac}} \left(\frac{\theta - \theta_w}{\theta_{\text{opt}} - \theta_w}\right), & \text{for } \theta_w < \theta < \theta_{\text{opt}} \\
(1 - \theta_{\text{fac}}), & \text{for } \theta \leq \theta_w
\end{cases}
\]

Here, \(\kappa_s\) is a constant representing the base respiration at 25°C per unit soil carbon.
Cs is the soil carbon in kgC m$^{-2}$. In the dynamic version of TRIFID, the soil carbon pool is increased by litterfall and decreased by flux to the atmosphere. In this experiment, vegetation dynamics are disabled and we solve for $C_s$ as a model parameter. $C_s$ is then held constant in time. Because there is only one soil layer, this parameter represents all types of soil carbon. $Q_{10RH}$ is a factor describing the sensitivity of heterotrophic respiration to leaf temperature. The moisture sensitivity function $f_\theta$ is characterized by an optimal value of soil water content, above and below which there is a linear decrease of respiration. $\theta_{fac}$ is a parameter which controls the strength of the linear sensitivity of respiration to $\theta$, $\theta_{opt}$ is the optimal values of soil water content at which maximum respiration occurs, and $\theta_w$ is the wilting soil water content, below which respiration is constant and does not decrease further. The soil moisture parameters $\Theta_{fac}$, $\theta_{opt}$ and $\theta_w$ are all optimized in this study.

Autotrophic respiration is split into maintenance and growth respiration. Maintenance respiration is given by the following equation:

$$R_m = R_{dc} \left\{ \frac{\beta + \left( \frac{N_r + N_s}{N_l} \right)}{N_l} \right\}$$  \hspace{1cm} (4.10)

where $R_{dc}$ is the canopy dark respiration and $\beta$ is the moisture dependence factor described in Eq. 4.8. $N_l$, $N_s$ and $N_r$ are the leaf, stem and root nitrogen contents, respectively. The considered version of TRIFID does not contain a dynamic nitrogen cycle. As a result, the nitrogen contents depend only on pool sizes (Cox, 2001). Growth respiration is described by:

$$R_g = r_g (GPP - R_m)$$  \hspace{1cm} (4.11)
rg is assumed to be equal for all plant functional types.

### 4.2.2.3 Phenology

The LAI of the canopy is modified by a variable $p$ that controls the phenological status of the vegetation. The balanced LAI $L_b$ represents the maximum or “full-leaf” value of LAI and is determined by allometric relationships to canopy height as a function of PFT. The variable $p$ ranges between 0 and 1 and is a multiplier of the balanced LAI $L_b$ ($\text{LAI} = pL_b$). $p$ varies the most in the spring and autumn to represent the effects of leafout and senescence. $p$ is controlled by the leaf mortality rate $\gamma$, which is a function of leaf temperature:

$$
\gamma = \gamma_0 \quad \text{for } T > T_{\text{off}}
$$

$$
\gamma = \gamma_0 \left[ 1 + 9\left( T_{\text{off}} - T \right) \right] \quad \text{for } T \leq T_{\text{off}}
$$

$T_{\text{off}}$ is a PFT-specific parameter below which leaf turnover increased by a factor of 10 when the leaf temperature drops 1°C below this value. $T_{\text{off}}$ for broadleaf trees is optimized in this study, while $T_{\text{off}}$ for needleleaf trees is set to the model default value of -30°C because temperatures at the five eddy covariance sites rarely reach values this low. $\gamma_0$ is the minimum value of leaf mortality and is set to a constant value of 0.25 yr$^{-1}$ in this version of TRIFFID for all PFTs. When the value of $\gamma$ exceeds twice the minimum value $\gamma_0$, leaves are dropped at a constant rate $\gamma_0$ and the value of $p$ is modified. When the value of $p$ is less than one and $\gamma$ is less than twice the minimum value $\gamma_0$, the value of $p$ is increased:
We modify the TRIFFID model structure to include a potentially improved representation of phenology. When using Eq. 4.13, cool temperatures in the spring (T < T_{off}) would often cause the LAI to drop after beginning to increase. Similarly, warm temperatures in autumn would often cause leaves to “grow back”, increasing LAI after senescence had begun. Therefore, we have added the additional constraint that \( \frac{dp}{dt} \) must be greater than or equal to zero in the first half of the year (before day of year 180). Similarly, \( \frac{dp}{dt} \) must be less than or equal to zero in the second half of the year (after day of year 180).

\[
\frac{dp}{dt} = -\gamma_p \quad \text{for } \gamma > 2\gamma_0 \\
\frac{dp}{dt} = \gamma_p (1 - p) \quad \text{for } \gamma \leq 2\gamma_0
\]

\[4.13\]

4.2.3 Model forcings

At each tower site, the TRIFFID model is forced with observed hourly meteorological data. Half-hourly data are averaged to hourly timesteps when necessary. Meteorological forcings include air temperature, precipitation, incoming shortwave radiation, wind speed and relative humidity. Incoming shortwave radiation was estimated by assuming a constant ratio of observed photosynthetically active radiation (PAR) to this quantity. Gap-filled meteorological datasets were used when available. Gap-filled air temperature and PAR datasets were available at most sites. A gap-filled precipitation product was available at Harvard Forest. Gaps in precipitation data were filled by using daily data from nearby surface meteorological stations downloaded from the National Climatic Data Center (NCDC). Gaps in other variables were filled using
monthly diurnal mean values. Because the focus of this study is interannual variability rather than long-term trends, we do not consider the possible effects of CO₂ fertilization in this study. Ambient CO₂ concentrations are set to a constant 370 ppm at all times at each site. In one of the assimilation experiments, observed soil moisture at the WLEF tower is used as a forcing rather than precipitation.

4.2.4 Model initialization

Model spinup to obtain the equilibrium soil carbon and forest cover for each site would take on the order of centuries and would therefore be computationally prohibitive to complete this data assimilation exercise. In addition, there is a considerable amount of error in the equilibrium vegetation cover obtained with TRIFFID as compared with IGPB data in eastern North America (Meissner et al., 2003). The distribution of plant functional types and forest canopy heights are therefore specified using values obtained from the literature (Table 4-1). Because there is only a single carbon pool, we specify the amount of carbon in this pool as a model parameter in each single-site optimization. In the joint optimization, we allow this parameter to be site-specific, adding four additional parameters. We consider only two PFTs: broadleaf and needleleaf trees. To limit the number of PFT-specific parameters in the optimization, the fractions of C₃ grass, C₄ grass and shrub are set to zero for all sites, although some small fraction (<5%) of the tower footprints may be occupied by these types. In all simulations, vegetation dynamics are disabled so that vegetation height, vegetation distribution, and soil carbon are unaffected by model fluxes and feedbacks.
4.2.5 Observational Constraints

Hourly values of observed net ecosystem exchange (NEE) are used to constrain TRIFFID model parameters in the optimization. Half-hourly observations are averaged to hourly when necessary, but gap-filled values of NEE are not used as model constraints. Fluxes computed during low-turbulence conditions are also screened using site-specific $u^*$ thresholds from the literature. In these simulations, we assume that errors in NEE are normally distributed and equal. The ramifications of these assumptions are discussed below.

4.2.6 Optimization Method

The goal of the optimization exercises is to maximize the likelihood $L$ of a set of TRIFFID model parameters $\theta_k$ given a set of tower observations $x$. If all observations are independently and identically distributed (IID) and follow a normal distribution, then the form of the likelihood function is:

$$L(x \mid \theta_k) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{1}{2} \left[ \frac{f(\theta_k, t_i) - x_i}{\sigma_i} \right]^2 \right)$$  \hspace{1cm} 4.14$$

The likelihood of a set of $n$ observations in $x$ is the product over the likelihoods of individual observations $x_1, x_2, ..., x_n$ at times $t_1, t_2, ..., t_n$. We assume that each observation is sampled independently from a Gaussian distribution with a mean $x_i$ and standard deviation $\sigma_i$ representing the variability due to observation noise and internal...
variability that is not captured by the model. Observation error is assumed to be constant in time, and for the joint optimization, constant across sites.

This formulation of the likelihood function follows previous studies that assume IID and normally distributed variability (Braswell et al., 2005; Hargreaves and Annan, 2002). There is some indication that random flux error may more closely follow a double exponential rather than a Gaussian distribution (Richardson et al., 2006), but we have chosen to implement a Gaussian framework for consistency with previous studies. The IID assumption is useful and reasonable approximation only if both observation errors and process noise are IID. Observational error related to turbulent sampling is likely to be independent because turbulent eddies in the atmospheric boundary layer are short-lived (Lenschow and Stankov, 1986; Berger et al., 2001) compared to the hourly flux measurement interval.

Because of the nonlinearity of the TRIFFID model, the parameter space is likely to be nonconvex, requiring the use of a global optimization algorithm. Based on previous benchmark results [Moles et al., 2004] we implement the SRES algorithm [Runarsson and Yao, 2000] to solve this system. SRES is a genetic optimization algorithm that uses a sequence of mutation and selection to improve an objective function. The algorithm starts with an initial random population of trial vectors that are evaluated by the objective function. Some fraction of well-performing members of this population are selected and used to generate new trial vectors (with random perturbation, akin to the effects of mutation in the evolution of natural populations). The repeated application of selection and mutation results in a generally improving objective function without the typical pitfalls of gradient based methods [Goldberg, 1989].
Finding the global optimum for this non-convex optimization problem is nontrivial. As with all global optimization problems without a known analytical solution, one cannot mathematically prove that the precise global optimum has been identified in a finite computation time. However, we implement an algorithm for detecting misconvergence that ensures a close approximation to the optimal solution. Specifically, we run the SRES algorithm for three sets of initial conditions and compare likelihood and parameter values at the end of each iteration. The two SRES runs with the highest likelihood are compared. If the parameters differ by less than 0.5% of the prior range, the solutions are, for practical purposes, identical and likely a good approximation to the global maximum.

One additional benefit of the SRES algorithm is the underlying genetic algorithm is implemented to run in parallel on a computer cluster. The evaluation time of TRIFFID is about 1 second of wall-time per ten site-years. One run of SRES requires on the order of $10^5$ model evaluations, or roughly 1-2 processor-days depending on the length of the site record. The joint optimization, which uses 37 site-years of data, requires about 7 processor-days. To limit run times to less than one day, we run each separate optimization on two processors in parallel, and the joint optimization is run on eight processors on the Penn State GEaRS high-performance computing cluster. Overall, the analysis shown in this study requires a total of about 35 processor-days of computations.
4.2.7 Experimental design

We run three assimilation experiments to test the hypotheses outlined in 4.1. First, the five sites are optimized separately using only the existing NEE observations at each site. For each site optimization, we optimize 22 model parameters, which are described in table 2. Second, all sites are optimized jointly. Soil carbon and leaf nitrogen are site-specific, adding 7 parameters for a total of 29 parameters in this optimization. Finally, we optimize the WLEF site using observed rather than modeled soil moisture. As described above, we run the SRES algorithm four times for each site optimization and compare the resulting parameters and values of the likelihood function. In addition to the optimizations, we also run the model once for each site using published model parameters for the purpose of comparing to the optimized model.

To test the hypothesis of nonconvexity in the objective function, we perform likelihood scans around the global solution for each model parameter. In a likelihood scan, one parameter is varied over the prior range while all other parameters are held constant at their optimal values. Several maxima in this likelihood scan indicate a nonconvex problem.

We test the second hypothesis by comparing optimized model NEE to model NEE calculated using published TRIFFID parameters (Cox, 2001). To avoid using gap-filled NEE in this comparison, we compute a mean seasonal cycle using only hours during which observations exist. The mean seasonal cycle consists of 18 20-day mean values. We note that taking a 20-day mean only for hours during which observations exist would cause a bias because more nighttime than daytime values are missing as a result of u-
screening. Therefore, we first calculate a mean for each hour over each 20-day average period to produce a mean diurnal cycle. These 24 hourly values are then averaged to estimate the 20-day mean NEE. The models are evaluated by how well the seasonal cycle matches the observations. R² values from the separately optimized model are compared to R² values from the joint optimization and from the unoptimized model using published parameters.

To evaluate how well the models reproduce interannual variability, we calculate “observed” annual NEE sums by summing all hours for each year over which observations exist. For missing hours, we substitute the monthly mean diurnal value calculated using all years in the site-record. These values are also substituted in the modeled timeseries. This technique eliminates the need to use gap-filled flux data, which are produced using other models (e.g. Falge et al., 2001) and may otherwise bias the comparison. This technique also avoids a false signal of interannual variability that may be caused by variations in the timing and length of data gaps from year to year. The models will be judged by the correlations (R² values) of simulated annual NEE sums with observed annual NEE sums.

Estimates of observed cross-site variability are made using the average annual NEE at each site. To avoid using gap-filled data, the annual NEE is computed by computing monthly diurnal mean values for every hour of each month across all years, then summing this value over all hours. Modeled annual NEE sums are computed with the same method using only hours during which observations exist.
4.3 Results and discussion

4.3.1 Nonconvexity of likelihood function

The considered data assimilation problem is nonconvex (Fig. 4-3). This example is from a likelihood scan of a parameter in the WLEF separate optimization. Several other nonconvexities were found in likelihood scans for other parameters and other site optimizations, but Fig. 4.3 is the clearest example. Nonconvexity in the full parameter space is also strongly implied in the convergence diagnostics in which different initial conditions for SRES result in different solutions (see below).

Nonconvexity of the parameter space was also found in other terrestrial carbon modeling data assimilation studies that used CO₂ concentrations rather than fluxes as data constraints. Rayner et al. (2005) and Vukicevic et al. (2001) both used gradient-based methods for parameter optimization and noted possible miconvergence in their solutions. Using an appropriate global optimization method is an important technical point. The typically applied local optimization methods such as tangent linear adjoint methods (Fennel et al., 2001; Schartau et al., 2001; Vukicevic et al., 2001) are prone to converge to just a local (as opposed to the global) minimum and hence suggest often biased results.

4.3.2 Convergence of SRES algorithm

For each optimization, three SRES runs were compared to assess whether they converged to the global solution. In all of the separate optimizations and in the joint
optimization, the values of the log likelihood function after 1,000 model iterations were within 0.5% of each other (Fig. 4-4). Despite this agreement, values of model parameters did not always match in all three SRES runs for any given optimization (Fig. 4-5). Parameters have not converged if the range of the three SRES runs is larger than 0.5% of the difference between the lower and upper bounds. The likelihood function is generally not sensitive to these parameters so that different values may yield solutions that are nearly equally likely. The parameters controlling the soil moisture of heterotrophic respiration ($\theta_{\text{opt}}$ and $\theta_{\text{fac}}$) do not converge at most sites; this is another strong indicator of the nonconvexity shown in Fig. 4-3. This also implies that some or all

![Figure 4-3: Scan of the log likelihood function over the prior range of the $\theta_{\text{opt}}$ (optimal soil water content for heterotrophic respiration) parameter for the WLEF individual optimization. Two maxima are visible, indicating a nonconvexity in this 1D scan. The use of gradient-based optimization algorithms may misconverge to a local solution rather than the desired global solution.](image)
of the SRES runs have misconverged to local rather than global solutions. A number of PFT-specific parameters do not converge when a PFT is a small fraction of the site vegetation. In these cases, the likelihood function is relatively insensitive to these parameters. For example, the broadleaf parameters \( n_{10}, T_{upp}, \) and \( T_{low} \) do not converge at Howland Forest, which is composed mainly of needleleaf trees. Conversely, the needleleaf versions of these parameters do not converge at UMBS, which only contains 2% needleleaf trees.

Although the values of parameters reported in this study produce reasonable fits to the observed data, we cannot guarantee that the SRES algorithm has converged to the global solution in any of the optimizations. Despite the apparent misconvergence of SRES in some cases, we contend that this algorithm still produces preferable results to gradient-based algorithms, which have been used in terrestrial data assimilation studies of nonconvex problems. Using a technique such as Markov Chain Monte Carlo (e.g. Braswell et al., 2005; Ricciuto et al., 2006; Sacks et al., 2006) would provide joint parameter probability density functions, allowing us to explore the issue of nonconvexity further. However, this is not feasible given the available computational resources. In all cases, we report parameters from the SRES run with the highest (least negative) log likelihood function. Parameters that have not converged when using the SRES algorithm are indicated in the analysis.
Because SRES is a minimization algorithm, we use the negative of the log likelihood as our objective function. In all five cases, the values of the objective function in runs 2 and 3 are within 0.5% of the value obtained in run 1. Results are similar for the joint optimization (not shown). In all optimizations, the run with the lowest value of the objective function is used to make model predictions.

Figure 4-4: Convergence of the SRES algorithm for each individual site optimization.
Figure 4-5: Convergence diagnostics of 4 model parameters the Howland Forest individual site optimization. In all three SRES runs, the soil carbon and $Q_{10H}$ parameters converge to nearly identical values. However, the soil moisture parameters $\theta_{\text{fac}}$ and $\theta_{\text{opt}}$, that control the soil moisture dependence of heterotrophic respiration, converge to different values in each of the three runs despite the similar likelihood values illustrated in Fig. 4-4.
4.3.3 Parameter estimates

The optimizations yield parameter estimates that differ considerably from published values in several cases (Table 4-3). Key differences in several parameters between the published and optimized parameters contribute to large discrepancies in predictions. In a few cases, the most likely parameter values hit the upper or lower bound. Some parameters did not converge to a single value, but instead varied considerably among the three SRES runs. This is indicative of nonconvexity in these parameters.

The heterotrophic respiration parameters (soil carbon CS, temperature sensitivity $Q_{10H}$, optimal moisture $\theta_{opt}$, and moisture dependence $\theta_{fac}$) generally did not converge in the separate optimizations. Although the optimized values of $Q_{10H}$ are reasonable and within the range of 1.3 to 3.3 expected based on the results of Raich and Schlesinger (1992), these values were not stable among SRES runs at Harvard Forest and Morgan Monroe. Values of the soil carbon parameter only converged at WLEF. The optimal moisture $\theta_{opt}$ was similar at all five sites, but did not converge at Howland Forest or Morgan Monroe. This parameter displayed the most nonconvexity in likelihood scans (Fig. 4-3). In three cases, the moisture factor $\theta_{fac}$ was edge-hitting, implying that heterotrophic sensitivity displays the maximum possible sensitivity to soil moisture. Values higher than one would cause negative respiration in dry conditions (Eq. 4.9).

The critical and wilting soil moisture parameters ($\theta_c$ and $\theta_w$ respectively) do not converge. This may be due to the relative lack of dry conditions in these eastern U.S. sites, yielding limited data to constrain these parameters that are associated
Table 4-3: Results of the parameter optimization for each separate optimization at HW (Howland), HV (Harvard), WL (WLEF), MM (Morgan Monroe) and UM (UMBS). Parameters that have converged are shown in bold, and are defined as values that are within 0.5% of each other in all three SRES runs and are not edge-hitting (upper and lower bounds and parameter units are shown in Table 4-2). For parameters that have not converged, results from the most likely of the three SRES runs are shown.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Published</th>
<th>HV</th>
<th>HW</th>
<th>WL</th>
<th>MM</th>
<th>UM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>10.0</td>
<td>4.81</td>
<td>4.827</td>
<td>9.81</td>
<td>10.7</td>
<td>7.37</td>
</tr>
<tr>
<td>Q(_{10H})</td>
<td>2.0</td>
<td>1.47</td>
<td>2.394</td>
<td>3.28</td>
<td>1.92</td>
<td>2.65</td>
</tr>
<tr>
<td>(\theta_{opt})</td>
<td>0.55</td>
<td>0.843</td>
<td>0.717</td>
<td>0.745</td>
<td>0.685</td>
<td>0.764</td>
</tr>
<tr>
<td>(\theta_{fac})</td>
<td>0.8</td>
<td>1.00</td>
<td>0.320</td>
<td>0.858</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(\theta_c)</td>
<td>0.30</td>
<td>0.247</td>
<td>0.129</td>
<td>0.240</td>
<td>0.285</td>
<td>0.239</td>
</tr>
<tr>
<td>(\theta_v)</td>
<td>0.13</td>
<td>0.00</td>
<td>0.0498</td>
<td>0.0534</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>n(_{0BL})</td>
<td>0.036</td>
<td>0.149</td>
<td>0.250</td>
<td>0.130</td>
<td>0.100</td>
<td>0.112</td>
</tr>
<tr>
<td>n(_{0NL})</td>
<td>0.030</td>
<td>0.140</td>
<td>0.045</td>
<td>0.0430</td>
<td>0.0500</td>
<td>0.0516</td>
</tr>
<tr>
<td>(\alpha_{BL})</td>
<td>0.06</td>
<td>0.0535</td>
<td>0.117</td>
<td>0.0314</td>
<td>0.0498</td>
<td>0.0353</td>
</tr>
<tr>
<td>(\alpha_{NL})</td>
<td>0.06</td>
<td>0.0274</td>
<td>0.021</td>
<td>0.483</td>
<td>0.249</td>
<td>0.340</td>
</tr>
<tr>
<td>Q(_{10VM})</td>
<td>2.0</td>
<td>2.23</td>
<td>1.74</td>
<td>2.36</td>
<td>2.15</td>
<td>2.04</td>
</tr>
<tr>
<td>T(_{lowBL})</td>
<td>-5.0</td>
<td>1.48</td>
<td>-40.0</td>
<td>-9.70</td>
<td>6.06</td>
<td>-1.83</td>
</tr>
<tr>
<td>T(_{lowNL})</td>
<td>-15.0</td>
<td>3.55</td>
<td>3.41</td>
<td>-1.80</td>
<td>10.0</td>
<td>-20.0</td>
</tr>
<tr>
<td>T(_{uppBL})</td>
<td>33.0</td>
<td>50.0</td>
<td>50.0</td>
<td>50.0</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>T(_{uppNL})</td>
<td>28.0</td>
<td>11.66</td>
<td>50.0</td>
<td>50.0</td>
<td>10.0</td>
<td>49.1</td>
</tr>
<tr>
<td>(D_c)</td>
<td>0.09</td>
<td>0.0192</td>
<td>0.0266</td>
<td>0.0249</td>
<td>0.0371</td>
<td>0.0282</td>
</tr>
<tr>
<td>R(_g)</td>
<td>0.25</td>
<td>0.107</td>
<td>0.0500</td>
<td>0.122</td>
<td>0.0500</td>
<td>0.0500</td>
</tr>
<tr>
<td>R(_{dc})</td>
<td>0.015</td>
<td>0.00822</td>
<td>0.0208</td>
<td>0.0183</td>
<td>0.0019</td>
<td>0.00866</td>
</tr>
<tr>
<td>Q(_{10RD})</td>
<td>2.0</td>
<td>1.49</td>
<td>1.61</td>
<td>2.00</td>
<td>1.100</td>
<td>1.33</td>
</tr>
<tr>
<td>T(_{off})</td>
<td>0.0</td>
<td>11.6</td>
<td>0.80</td>
<td>9.30</td>
<td>14.3</td>
<td>11.4</td>
</tr>
<tr>
<td>(\beta_{off})</td>
<td>0.90</td>
<td>0.678</td>
<td>0.850</td>
<td>0.672</td>
<td>0.172</td>
<td>0.517</td>
</tr>
<tr>
<td>(\gamma_p)</td>
<td>14.5</td>
<td>18.2</td>
<td>13.4</td>
<td>16.5</td>
<td>20.0</td>
<td>38.7</td>
</tr>
<tr>
<td>-(Log L)</td>
<td>N/A</td>
<td>157209</td>
<td>100989</td>
<td>69511</td>
<td>68554</td>
<td>61217</td>
</tr>
</tbody>
</table>
with the limitation of photosynthesis during dry conditions. The most likely values of critical SWC $\theta_c$ are similar at the four sites dominated by deciduous forests (excluding Howland forest) despite failing to converge. The wilting SWC $\theta_w$ is zero at three of the sites (edge-hitting), and does not converge at the other two sites.

Top-leaf nitrogen ($n_{l0}$) and quantum efficiency ($\alpha$) converge for the dominant PFT at each site. At the sites dominated by deciduous forests, the optimized $n_{l0BL}$ averages about three times the published value, indicating that modeled values of $V_{\text{max}}$ were too small (4.2.2). At Howland forest, which is dominated by needleleaf forests, the value of $n_{l0NL}$ increased by about 50% over the published value. $n_{l0NL}$ also converges at Harvard, WLEF and UMBS. WLEF and UMBS values are similar to Howland, but at Harvard $n_{l0NL}$ is about a factor of three larger. Optimized values of $\alpha$ are close to or smaller than the published values. Harvard Forest has the largest value of $\alpha_{BL}$ among the deciduous sites, followed by Morgan Monroe, UMBS and WLEF. At Howland, the value of $\alpha_{NL}$ is a factor of three smaller than the published value, which is slightly smaller than the optimized value at Harvard and much smaller than at WLEF.

Other parameters related to photosynthesis generally converge. $Q_{10VM}$, the temperature sensitivity of $V_m$, (Eq. 4.3) is near the published value of 2.0 in all cases. The critical humidity deficit $D_c$ is well-constrained at all sites, but is a factor of three to five smaller than the published value in all cases. This implies that photosynthesis is more sensitive to higher vapor pressure deficits (VPDs) at the eddy covariance sites than implied by the published TRIFFID model (Eq. 4.2). The temperature sensitivity of dark respiration, $Q_{10RD}$, converges at all sites but does not have as tight of a range as $Q_{10VM}$. 
$T_{\text{upp}}$ and $T_{\text{low}}$ do not converge at all sites. $T_{\text{lowBL}}$ converges at the deciduous-dominated sites, but varies widely over a range of $15^\circ$C. We hypothesize that $T_{\text{off}}$, the phenological leaf-off parameter, is a stronger control over the low temperature limitation of broadleaf photosynthesis. $T_{\text{lowNL}}$ converges at Harvard, Howland and WLEF. Values of $T_{\text{lowNL}}$ are similar at Harvard and Howland and slightly lower at WLEF; all three values are higher than the published value of $-15^\circ$C. $T_{\text{uppBL}}$ hits the upper bound in all five optimizations, suggesting that leaf temperatures are never warm enough to limit photosynthesis at these sites.

The phenological parameters $T_{\text{off}}$ and $\gamma_p$ converges while the moisture parameter $\beta_{\text{off}}$ does not. $T_{\text{off}}$ ranges at the deciduous-dominated sites between 9.3 at WLEF to 14.3 $^\circ$C at Morgan Monroe. $T_{\text{off}}$ appears to be a function of mean annual site temperature, with WLEF being the coldest and Morgan Monroe being the warmest of the optimized sites. This is probably indicative of species that are more tolerant to cold at WLEF. At Howland forest, the value of $T_{\text{off}}$ is the smallest at 0.8$^\circ$C, but this site has a small deciduous component. Aside from Howland, the optimized values of $T_{\text{off}}$ are considerably higher than the published value of 0.0$^\circ$C, indicating that the modeled growing season is too long when using published parameters. Values of $\gamma_p$ at the four deciduous-dominated sites are slightly less than the published value of 20 yr$^{-1}$, ranging from 13.4 yr$^{-1}$ at Morgan Monroe to 18.2 yr$^{-1}$ at WLEF. A higher value of $\gamma_p$ indicates a more rapid leaf-on and leaf-off. This parameter is inversely related to mean annual site temperature, implying that deciduous leaf-on and leaf-off are faster at WLEF than at Morgan Monroe.
Estimated parameters from the joint optimization were generally well-constrained but did not agree in all cases with the results of the individual optimizations (Table 4-4). We would expect the jointly optimized parameters to represent roughly the mean of the parameters from the separate optimizations. In the joint optimization, soil carbon (CS) and top-leaf nitrogen (n_{l0}) are site-specific parameters, adding seven parameters to the analysis. To avoid having two site-specific nitrogen parameters, we did not treat n_{l0} as a PFT-specific parameter in the joint optimization. The other 20 parameters are assumed to be equal for all sites. This analysis implicitly assumes that variations in CO_{2} fluxes across sites are caused solely by differences in soil carbon, leaf nitrogen content, PFT distribution, canopy height and climate forcing.

In the joint optimization, heterotrophic respiration parameters are generally poorly constrained or unrealistic. The soil carbon varies widely from site to site as compared to the results of the separate optimization experiments. In fact, the optimized soil carbon content is zero at both Harvard Forest and Morgan Monroe, and as high as 12 kg m^{-2} at WLEF. A soil carbon content of zero implies that there is no heterotrophic respiration at the sites, which is highly unrealistic. This could be due to an overestimation of the autotrophic respiration component, poorly modeled wintertime respiration, or biases in the nighttime eddy covariance observations. Because these values do not agree with those obtained in the separate optimizations, it is likely that the model has failed to capture all of the processes controlling intersite variability in heterotrophic respiration. The value of Q_{10H} is 4.0, which is the upper bound used in the optimization. Values of Q_{10} exceeding 4.0 have been reported when fitting nighttime eddy covariance flux data to
soil temperature (e.g. Van Dijk and Dolman, 2004), but we suspect this value is unrealistic because it is larger than any $Q_{10H}$ obtained in the individual optimizations.

Other parameters converge and are closer to those obtained in the separate optimizations. Values of $\theta_{\text{opt}}$ and $\theta_{\text{fac}}$ are comparable to the published values, although $\theta_{\text{opt}}$ is smaller than values obtained in the separate optimizations. The critical SWC $\theta_c$ is similar to the values obtained in the separate optimizations. The site-specific values of leaf nitrogen $n_{l0}$ converge and are a factor of two to three larger than the published values, again largely in agreement with the separate optimizations. Most of the

Table 4-4: Results of the joint parameter optimization, which uses observations from all sites as a data constraint. Parameters that have converged and that are not edge-hitting are shown in bold, (upper and lower bounds and parameter units are shown in Table 4-2)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Published</th>
<th>Optimized</th>
<th>Parameter</th>
<th>Published</th>
<th>Optimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS (HV)</td>
<td>10.0</td>
<td>0.00</td>
<td>$\alpha_{\text{BL}}$</td>
<td>0.06</td>
<td>0.0446</td>
</tr>
<tr>
<td>CS (HW)</td>
<td>10.0</td>
<td>5.59</td>
<td>$\alpha_{\text{NL}}$</td>
<td>0.06</td>
<td>0.0239</td>
</tr>
<tr>
<td>CS (WL)</td>
<td>10.0</td>
<td>12.0</td>
<td>$Q_{10VM}$</td>
<td>2.0</td>
<td>2.26</td>
</tr>
<tr>
<td>CS (MM)</td>
<td>10.0</td>
<td>0.00</td>
<td>$T_{\text{lowBL}}$</td>
<td>-5.0</td>
<td>-1.05</td>
</tr>
<tr>
<td>CS (UM)</td>
<td>10.0</td>
<td>2.97</td>
<td>$T_{\text{lowNL}}$</td>
<td>-15.0</td>
<td>50.0</td>
</tr>
<tr>
<td>$Q_{10H}$</td>
<td>2.0</td>
<td>4.00</td>
<td>$T_{\text{uppBL}}$</td>
<td>33.0</td>
<td>1.475</td>
</tr>
<tr>
<td>$\theta_{\text{opt}}$</td>
<td>0.55</td>
<td>0.447</td>
<td>$T_{\text{uppNL}}$</td>
<td>28.0</td>
<td>30.0</td>
</tr>
<tr>
<td>$\theta_{\text{fac}}$</td>
<td>0.80</td>
<td>0.980</td>
<td>$D_c$</td>
<td>0.09</td>
<td>0.0246</td>
</tr>
<tr>
<td>$\theta_c$</td>
<td>0.30</td>
<td>0.216</td>
<td>$R_p$</td>
<td>0.25</td>
<td>0.005</td>
</tr>
<tr>
<td>$\theta_q$</td>
<td>0.13</td>
<td>0.00</td>
<td>$R_{dc}$</td>
<td>0.015</td>
<td>0.0143</td>
</tr>
<tr>
<td>$n_{l0}(HV)$</td>
<td>0.036</td>
<td>0.112</td>
<td>$Q_{10RD}$</td>
<td>2.0</td>
<td>1.515</td>
</tr>
<tr>
<td>$n_{l0}(HW)$</td>
<td>0.030</td>
<td>0.087</td>
<td>$T_{\text{off}}$</td>
<td>0.0</td>
<td>11.5</td>
</tr>
<tr>
<td>$n_{l0}(WL)$</td>
<td>0.036</td>
<td>0.080</td>
<td>$\beta_{\text{off}}$</td>
<td>0.90</td>
<td>0.00</td>
</tr>
<tr>
<td>$n_{l0}(MM)$</td>
<td>0.036</td>
<td>0.083</td>
<td>$\gamma_p$</td>
<td>20.0</td>
<td>15.8</td>
</tr>
<tr>
<td>$n_{l0}(UM)$</td>
<td>0.036</td>
<td>0.113</td>
<td>$-(\log L)$</td>
<td>N/A</td>
<td>465957</td>
</tr>
</tbody>
</table>
remaining parameters converge and are within the range of values reported in the separate optimizations. The exceptions occur with parameters that do not converge in the separate optimizations. $T_{\text{uppNL}}$, which does not converge in the separate optimizations, is 30°C; this is near the published value. Further implications of our parameter fits on modeled fluxes are discussed below.

4.3.4 Variability from diurnal to seasonal timescales

The optimized version of TRIFFID reproduces 20-day average model fluxes much better than the version with published parameters (Fig. 4-6); the primary reason is that the amplitude of the seasonal cycle simulated using the published version of TRIFFID is too small. We note that this version of TRIFFID, as used in coupled carbon cycle/GCM studies (Cox et al., 2000; Friedlingstein et al., 2006; Matthews et al., 2005a; Meissner et al., 2003), would be spun up to equilibrium such that the soil carbon changes from its published value of 10 kg m$^{-2}$. We also note, however, that the published value of $n_{T0}$ is roughly a factor of three smaller than the optimized value. We hypothesize that this parameter, rather than soil carbon, is the primary reason for the low-amplitude seasonal cycle; there is not enough daytime uptake. This effect is illustrated in the mean growing season diurnal cycle at all sites (Fig. 4-7). Estimates of nighttime fluxes are reasonable, which implies that NPP is underestimated rather than $R_{HI}$ that is overestimated. In addition, multiplying the values of $n_{T0}$ by a factor of three while holding other parameters at their published values produces a diurnal cycle remarkably similar to that of the optimized models (not shown); therefore, we conclude that an underestimate of $V_{\text{max}}$ is
the biggest source of the discrepancies between the models in both figures 4-7 and 4-8. Optimized values of $V_{\text{max}}$ also more closely agree with the results of other data assimilation experiments (Knorr and Kattge, 2005; Rayner et al., 2005).

Figure 4-6: Mean seasonal cycle of NEE over the available tower record for the observed NEE, model NEE using published parameters, and model NEE using the optimized parameters from the separate optimizations. Each point represents a 20-day average of hourly NEE values. Using published parameters results in a poorly-timed growing season and not enough growing season uptake. The optimized models reproduce the phase and amplitude of the seasonal cycle well, with the exceptions of Howland and WLEF, where too much uptake is modeled in the latter part of the growing season.

In addition to the poor representation of the seasonal and diurnal amplitudes by the published model, there are also timing problems with the seasonal cycle. The growing season begins too early, especially at Harvard and UMBS. This is a result of the $T_{\text{off}}$ parameter that is too low, allowing leaf on to begin up to three weeks early. When
using published parameters, respiration overtakes photosynthesis relatively early in the growing season as a result of increasing soil temperature, and carbon efflux peaks in early September. The observations and optimized models generally have maximum efflux in October.

Slight degradation of the seasonal fits occurs when the parameters from the joint optimization are used (Fig. 4-8). The separate and joint optimizations produce nearly identical seasonal cycles, but the timing of spring uptake is too early for the joint optimization, most notably at Harvard and UMBS. The amplitude of the seasonal

Figure 4-7: Mean diurnal cycle of NEE over the growing season (defined as June-August) averaged over all years in the available tower record. Shown are observed NEE (x), model NEE using published parameters (dashed), and model NEE using the optimized parameters from the separate optimizations (solid). Using published parameters results in a diurnal cycle that is too small in magnitude. In most cases, estimates of nighttime CO$_2$ flux (respiration) are near the observed, but estimated daytime fluxes are too small. The optimizations improve this considerably, although still slightly underestimating the amplitude at Harvard, WLEF and UMBS.
cycle is also slightly too large at WLEF.

Figure 4-8: As in Figure 4-6, except showing a comparison of the separate and joint optimizations to observations. The joint and separate optimization produce similar results in all cases, with a slight degradation in the timing of the seasonal cycle when using the parameters from the joint optimization.

4.3.5 Interannual variability

Although diurnal and seasonal variability are well-captured by the optimized models, interannual variability is poorly simulated (Fig. 4-9). Our optimization of model parameters sometimes leaves the skill of modeling interannual variability virtually unchanged. At Howland forest, $R^2$ is near zero before and after optimization. The correlation between observed and modeled interannual variability does improve at
Harvard forest from 0.00 to 0.29 for the separate optimization, and to 0.35 for the joint optimization. At WLEF, the correlation improves from 0.00 to 0.16 for the separate optimization and 0.29 for the joint optimization. The optimized correlations at both sites are not statistically significant (p < 0.05). At Morgan Monroe, the separate optimization improves $R^2$ slightly but the joint assimilation does not. Curiously, the highest $R^2$ is obtained at UMBS when using the published parameters ($R^2 = 0.86$). The correlations degrades when using parameters from the separate optimization ($R^2 = 0.54$) and the joint optimization ($R^2 = 0.67$). However, the published parameters yield highly biased estimates of site-mean NEE at

![Graphs showing comparisons of observed annual NEE sums to simulated NEE sums computed using published parameters, parameters from the separate optimizations, and parameters from the joint optimization. Despite the ability of the model to capture diurnal and seasonal cycles, TRIFFID does not capture interannual variability at any tower in a statistically meaningful way (p < 0.05).]

**Figure 4-9:** Comparisons of observed annual NEE sums to simulated NEE sums computed using published parameters, parameters from the separate optimizations and parameters from the joint optimization. Despite the ability of the model to capture diurnal and seasonal cycles, TRIFFID does not capture interannual variability at any tower in a statistically meaningful way (p < 0.05).
UMBS and the other four towers. This version of the model predicts that all five towers are sources of CO₂ to the atmosphere, which is only true at WLEF. As discussed above, the primary reason for this bias is that daytime update is underestimated because values of \(V_{\text{max}}\) are too small.

Our analysis suggests that key processes controlling interannual variability at these sites are missing in the TRIFFID model. This results is somewhat surprising because of the relative success of the coupled TRIFFID/Hadley model in reproducing variability in the observed global CO₂ growth rate (Jones et al., 2001). That study focused on ENSO variability, which may not be a factor in governing interannual variability of NEE in the eastern U.S. One hypothesis to explain the apparent lack of skill of the TRIFFID model to reproduce observed interannual variability is that this variability in an eddy covariance tower footprint (~1km²) is highly localized and difficult to capture by a large-scale model. Spatial coherence of tower-measured NEE variations related to climate anomalies has been documented (e.g. Butler et al., in preparation; Ciais et al., 2005), but not in terms of interannual variability of annual NEE sums.

Our version of TRIFFID does not account for several important mechanisms that may be important in governing interannual variability. Because of the single soil pool and lack of a nitrogen cycle, TRIFFID may overlook the important effect of lagged nutrient availability (Vukicevic et al., 2001). TRIFFID fails to capture interannual variability at WLEF, despite the relative success of a simpler model (Ricciuto et al., 2006) when the same observations are assimilated. One of the primary drivers of interannual variability in that study was soil moisture, suggesting that the simple 1-D bucket model in TRIFFID may be insufficient to reproduce the actual site hydrology.
Because soil moisture observations are available at WLEF, we test this hypothesis by driving TRIFFID with the observed SWC rather than modeling this variable (Fig 4-10).

Although the simulated interannual variability agrees slightly better with observations when using site hydrology, this improvement is still relatively small. Poor simulations of interannual variability as measured as eddy covariance are not unique to TRIFFID. SipNET also fails to capture observed interannual variability at Harvard and Howland forests (Braswell et al., 2005; Sacks et al., 2006). Finding a successful process-based model to simulate interannual variability at an eddy covariance site remains an
open challenge. This failure to capture important features of the interannual variability of CO₂ fluxes casts doubt on the skill of century-scale predictions.

4.3.6 Intersite variability

The joint optimization is able to capture differences in site-average annual NEE sums when we allow \( n₀ \) and soil carbon to vary as site-specific parameters (Figure 4-11). Differences in NEE are caused by differences in these parameters, PFT distribution, canopy height and climate forcing. Although this result is encouraging, versions of TRIFFID that are used for prediction must spin the soil carbon up to equilibrium. It may be inappropriate to compare gridscale estimates of soil carbon to site-specific estimates because of large variability over small spatial scales. This also suggests that a single PFT-specific leaf nitrogen parameter is insufficient. Promising avenues to improve the skill of the TRIFFID model would be to add a dynamic nitrogen cycle and/or add more PFTs to adequately capture inter-site variability in leaf nitrogen.

4.4 Conclusion

We estimate model parameters for the TRIFFID dynamic vegetation model using eddy covariance data from five tower sites, both separately and jointly. The published parameter values underestimate photosynthesis and result in poor phasing of the seasonal cycle at all sites. The parameter space in the model is nonconvex, requiring the use of
global optimization algorithm. We use SRES, an evolutionary algorithm, to perform this optimization. Estimating the parameter values vastly improved model hindcasts of diurnal and seasonal cycles. The leaf nitrogen parameter increased by a factor of 2-3 at all sites, improving estimates of net CO₂ uptake. Additionally, the phenological parameter describing the critical leaf-off temperature (T_{off}) increased by 10°C, resulting

Figure 4-11: Comparison of modeled and observed NEE across sites. Modeled NEE is from the joint assimilation experiment in which one set of parameters is used to model all sites with the exception of soil carbon and leaf nitrogen, which are site-specific parameters. The joint assimilation captures inter-site variability, but fails to capture interannual variability at any site.
in better phasing of the seasonal cycle. Some parameters did not converge, indicating shortcomings in the model structure, biases in the observations, or insufficient observational constraints. Interannual is poorly simulated by both the published and the estimated parameters. This indicates that mechanisms are missing in the model structure. Cross-site variability is simulated well, indicating that soil carbon, leaf nitrogen, canopy structure and climate forcing are largely responsible for differences among sites.
Chapter 5

Conclusion

5.1 Summary

Data assimilation techniques were used to assimilate carbon cycle observations at a range of temporal and spatial scales. First, we used the Markov Chain Monte Carlo (MCMC) technique to estimate carbon cycle parameters in a global model. The residuals of the model displayed significant temporal autocorrelation. We accounted for this autocorrelation by fitting an additional model parameter. We found that neglecting temporal autocorrelation caused biased and overconfident parameter PDFs and predictions. The probability density functions (PDFs) of model parameters were not normally distributed and displayed significant correlations in some cases. Using an observed, rather than modeled, time series of global-average temperatures resulted in better constrained parameters and predictions. This occurred because observed CO₂ growth rate anomalies combined with observed interannual temperature variations constrained the respiration temperature sensitivity (Q₁₀) parameter. This constraint on respiration in turn constrained the carbon fertilization parameter (β).

We found that, under the S550 stabilization scenario, the uncertainty in predicted allowable CO₂ emissions increased with time. Given the adopted model structure, the future terrestrial sink was more uncertain than the oceanic sink. Assimilating hypothetical observations of annual terrestrial and oceanic CO₂ fluxes with similar uncertainties considerably reduced sink uncertainties. Terrestrial flux observations had
slightly more impact in reducing sink uncertainty than oceanic flux observations. Because observational estimates of ocean fluxes exist, this study stresses the need for independent estimates of global-scale terrestrial fluxes. These estimates will help to constrain parameters related to terrestrial carbon cycle feedbacks, which are currently very uncertain. Continued monitoring of the carbon cycle will also enable the design of better carbon cycle models and will assist in the design of possible carbon management strategies.

We then shifted focus from a global scale to an ecosystem scale by analyzing seven years of WLEF net ecosystem exchange (NEE) observations from the period 1997-2004. Systematic errors in these NEE sums were corrected, and random errors related to turbulent variability and uncertainty in gap-filling were estimated. Interannual variability of annual NEE at WLEF was found to be statistically significant with respect to the estimated random error. Climate was found to have an important impact on the variability of seasonal and annual NEE sums. Standard gap-filling models were extended to include soil moisture and soil temperature to produce a model of NEE variability across all timescales from diurnal to interannual at WLEF. We used the Markov Chain Monte Carlo technique in order to assimilate hourly observations of NEE into this model and to estimate probability density functions of all model parameters. Parameter PDFs were then used to estimate NEE and uncertainties, which we compared to observations.

The choice of u* threshold did not affect patterns of interannual variability except at very high or low values. WLEF was a source of carbon to the atmosphere from 1997-2004 regardless of the choice of u* threshold. The range of interannual variability was greater in the spring and summer than in autumn and winter, and greater during the day
than at night. Spring variability was related primarily to soil temperature, which was correlated to the date of leafout. Summer variability may be caused by both temperature and moisture. The competing effects of temperature and moisture tend to limit the magnitude of autumn variability, and wintertime variability is small despite large variations in temperature because photosynthesis and respiration are minimal during dormant conditions.

Estimated interannual model parameters were generally not constrained by the prior bounds. Parameter PDFs were not all Gaussian, and these PDFs changed depending on whether daytime, nighttime or all NEE data are used in the assimilation. This reflects significant differences between estimated nighttime and daytime respiration for the same given soil moisture and temperature. The interannual model reproduced interannual variability in annual and growing season NEE with moderate success. The model was unsuccessful in reproducing interannual variability in the transition seasons of spring and autumn. The model underestimates the range of interannual variability.

Finally, TRIFFID model parameters were optimized using eddy covariance data from five tower sites. Optimizations were performed separately for each site and jointly, using all sites simultaneously. The published TRIFFID parameters underestimated photosynthesis and had a poor phasing of the seasonal cycle at all sites. We found nonconvexities in the parameter space, requiring the use of global optimization algorithm to perform the data assimilation. We used an evolutionary algorithm to perform this optimization. Model parameters did not always converge in the SRES optimizations. However, optimizing TRIFFID vastly improved model predictions of diurnal and seasonal cycles. The leaf nitrogen parameter increased by a factor of 2-3 at all sites,
improving estimates of net uptake. Additionally, the phenological parameter $T_{\text{off}}$ was raised by 10°C, causing better phasing of the seasonal cycle. The misconvergence of some parameters indicated shortcomings in the model structure or biases in the observations. Cross-site variability was simulated well, indicating that soil carbon, leaf nitrogen, canopy structure and climate forcing are largely responsible for differences among sites. Interannual variability was not simulated well by either the published or optimized version of TRIFFID, indicating that key processes were missing in the model structure.
Bibliography


Kaminski, T., Knorr, W., Rayner, P.J. and Heimann, M., 2002. Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle. Global Biogeochem Cy, 16(4), -.


fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models. Tellus B, 51(2), 343-366.


estimates of evapotranspiration in a northern Wisconsin forest. Global Change Biol, 8(12), 1253-1265.


Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T.,


VITA

Daniel Ricciuto

Daniel Ricciuto was born in Stoughton, Massachusetts in 1978. He grew up in Holbrook, Massachusetts and Wilmington, Delaware, and graduated from Brandywine High School in 1996. He attended Duke University in Durham, North Carolina, where he obtained a Bachelor of Science degree in Physics and a minor in German in 2000. Following graduation, Daniel began graduate study at The Pennsylvania State University in the Department of Meteorology under the direction of Dr. Kenneth J. Davis. During this time, he acquired familiarity with eddy covariance data processing and data assimilation techniques, and developed a dissertation topic that combined these subjects. He also became a member of the Meteorology Honor Society, Chi Epsilon Pi, and served as an officer of this organization from 2002-2004. Daniel has been married to Kristin Lala Ricciuto since August of 2004.