DISSERTATION

CHARACTERIZING UNCERTAINTY FOR REGIONAL CARBON CYCLE MODELING

Submitted by

Lara Prihodko

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2004

COLORADO STATE UNIVERSITY

June 1, 2004

WE HEARBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY LARA PRIHODKO ENTITLED CHARACTERIZING UNCERTAINTY FOR REGIONAL CARBON CYCLE MODELING BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Roger A. Pielke Sr.

Dennis S. Ojima

Gregory P. Asner

A. Scott Denning

Adviser

Dan Binkley

Department Head/ Director

ABSTRACT OF DISSERTATION

CHARACTERIZING UNCERTAINTY FOR REGIONAL CARBON CYCLE MODELING

The discrepancy between current estimates of carbon dioxide emissions from natural processes, disturbance of the land surface and fossil fuel combustion, and the magnitude of accumulated CO_2 in the atmosphere as measured by the global flask network indicates that there must be a terrestrial and/or oceanic repository for the unaccounted carbon. The imbalance in estimates has become a major focus of research during the last two decades, particularly since our ability to predict and manage CO_2 concentrations in the future requires a thorough understanding of this disparity. A major challenge of current carbon cycle research is to reconcile the differences between top-down (eg. global inversions) and bottom-up (eg. process modeling, inventories) approaches. The research presented in this dissertation addresses some of the key aspects of model-data fusion that will be important within the framework of continental scaling efforts such as the North American Carbon Program.

The geographical locus of the research is the WLEF-TV tower, a very tall tower located in the Chequamegon National Forest near Park Falls, Wisconsin. Micrometeorological and eddy covariance measurements have been made since 1995 at three levels (30, 122, 396 meters) and many additional studies are underway in the region, forming the Chequamegon Ecosystem-Atmosphere Study (ChEAS). The vegetation of the region is heterogeneous, composed of upland Jack and Red Pine, mixed hardwoods and lowland conifers.

In a series of analyses, sources of error, potential for bias and impacts of model and data choices in the context of regional carbon cycle modeling of temperate, forested ecosystems were evaluated. A comprehensive sensitivity analysis was conducted using a Monte-Carlo style framework to evaluate the sensitivity of a commonly used, complex biophysical land surface model (Simple Biosphere Model, v.2; SiB2) to its parameterization through time. The results of the sensitivity analysis were then used within a Bayesian framework to calculate the uncertainty of land surface fluxes predicted by the model attributable to the parameterization. Finally the potential for uncertainty due to the spatial representation of the land surface was evaluated with a version of the SiB2 model coupled to the Regional Atmospheric Modeling System (RAMS). The results of this dissertation suggest that: there is a quantifiable level of model-data mismatch that could be used as a prior uncertainty in regional atmospheric inversions, compensation among parameters in complex biophysical land surface models complicates model optimization, uncertainty due to model parameterization was greater than the uncertainty attributable to mischaracterization of land surface heterogeneity and the potential for reduction in uncertainty appears to be greatest during the periods of rapid land surface change surrounding leaf-on in the spring and leaf-off in the autumn.

Lara Prihodko
Graduate Degree Program in Ecology
Colorado State University
Fort Collins, CO 80523
Summer 2004

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. A Scott Denning, for his patience, support and guidance through the years. His encouragement, enthusiasm and belief in me as a scientist has made all the difference. Thank you also to my other committee members: Dr. Roger Pielke Sr., Dr. Dennis Ojima and Dr. Gregory P. Asner for their support and insight. And a special thank you to Dr. Mike Lefsky, who came to my aid in a crisis.

I would also like to thank Dr. Ken Davis and Dr. Peter Bakwin for providing the data from the WLEF tall tower site in Park Falls Wisconsin, without which this research could not have been completed. And thank you as well to Dr. Mel Nicholls for his assistance with the coupled SiB-RAMS model.

I would also like to thank the entire Denning research group for their camaraderie. In particular, Connie Uliasz who makes everything run smoothly, John Kleist our computer guru and Ian Baker and Neil Suits for their patience in answering questions and bouncing around ideas.

I would also like to extend a special thank you to my friends for their tireless ears, general support and relief from the daily grind.

And finally, I am indebted to my family, especially my husband Niall Hanan, son Desmond and mother Stella, for their unfailing support. This would have been impossible without your infinite patience. You were there in crisis, exhaustion and exhilaration! Thank you cannot express my gratitude.

This work was supported by a NASA Earth System Science Fellowship – NGT5-30174 and NGT5-30182 and National Science Foundation Grant DEB-9977066.

DEDICATION

to my family for their loving support

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

Global atmospheric carbon dioxide concentrations ([CO_2]) have been increasing since the Industrial Revolution as a result of anthropogenic emissions. However, the rate at which measured atmospheric [CO_2] is increasing is inconsistent with these emissions, it is less than expected. This discrepancy between current estimates of carbon dioxide emissions from natural processes, disturbance of the land surface and fossil fuel combustion, and the magnitude of accumulated CO_2 in the atmosphere as measured by the global flask network (http://www.cmdl.noaa.gov/ccgg/index.html) indicates that there must be a terrestrial and/or oceanic repository for the unaccounted carbon (Houghton *et al.* 1990; Schimel 1995; Tans and White 1998). The imbalance in estimates has become a major focus of research during the last two decades, particularly since our ability to predict and manage CO_2 concentrations in the future requires a thorough understanding of this disparity.

The current understanding is that about half of the unaccounted for carbon is in the terrestrial biosphere of the Northern hemisphere. This conclusion was drawn from the results of atmospheric inversions which use seasonal and interannual patterns of [CO_2] to infer sources and sinks of CO_2 in the global biosphere (Tans *et al.* 1990; Bousquet *et al.* 2000; Gurney *et al.* 2002; Rodenbeck *et al.* 2003). Coupled biosphere-atmosphere models as well as observations from the global flask network are used in these inversions to locate areas with source and/or sink activity. The stable isotope composition of CO_2 contributes additional information to atmospheric inversions and is used to distinguish oceanic from terrestrial fluxes (Ciais *et al.* 1995; Rayner *et al.* 1999). Recent land based modeling and inventory estimates designed to reduce uncertainty in the calculation of sources and sinks of CO_2 in the coterminous United States suggest a land carbon sink ranging from 0.08 Pg C yr^{-1} to 0.58 Pg C yr^{-1} (Schimel *et al.* 2000; Pacala *et al.* 2001) consistent with atmospheric inversion studies (Gurney *et al.* 2002).

There is, however, persistent uncertainty in the precise magnitude and location of the sink(s). This uncertainty is the result of an incomplete understanding of how to observe, model and predict the appropriate processes in space and time. It is also the result of a limited capacity, both physical and computational, to observe and model what is understood. The practical consequences are global atmospheric inversions that are poorly constrained spatially due to a paucity of observations in certain geographic areas and due to the spatial resolution at which they can be run. Land based modeling and inventory estimates of sources and sinks of CO_2 are equally limited by a lack of available and appropriate data. Theoretically, uncertainties stem from simplified assumptions, such as those concerning the seasonal and interannual variability of isotope discrimination within terrestrial ecosystems (Enting *et al.* 1995; Rayner *et al.* 1999) and from incomplete understanding, such as how interannual variability in photosynthesis and respiration might result from disturbance and climatic variations (Bousquet *et al.* 2000).

There is also a lack of understanding of how land surface-atmosphere interactions and land surface processes scale through space and time (Oechel *et al.* 2000;

Vourlitis *et al.* 2000; Kaminski *et al.* 2001; Rahman *et al.* 2001; Geels *et al.* 2004). Atmospheric inversion models must specify time-space patterns of flux variations *a priori*. Errors in these patterns are inevitably aliased into errors in the estimated fluxes from these regions (Kaminski *et al.* 2001). Thus, an understanding of the uncertainty in the estimates of these space-time flux patterns and the propagation of this uncertainty through the inversion is critical (Engelen *et al.* 2002). Otherwise an accurate estimate of the uncertainty in the fluxes calculated from the inversions is impossible. Further, any improvement in these space-time patterns will reduce the uncertainty in the *a posteriori* fluxes.

Over the last ten years, large-scale field experiments (e.g. FIFE (Hall and Sellers 1995); HAPEX (Prince *et al.* 1995); BOREAS (Sellers *et al.* 1995a); LBA (Avissar *et al.* 2002)) and the development of international observational networks (FLUXNET; Baldocchi *et al.* 2001) have increased our understanding of regional flux patterns considerably, though the observations are typically limited in either space or time. The United States, Canada and Mexico have recently embarked on an ambitious, multi-year program (the North American Carbon Program; NACP) to better understand the role North American ecosystems play in the global carbon cycle and to reconcile the differences between top-down (eg. global inversions) and bottom-up (eg. process modeling, inventories) approaches (Wofsy and Harriss 2002; Denning *et al.* 2004). The program consists of long and short-term measurements from plant to continental scale. It will include process level studies, inventories and atmospheric concentrations for the land portion. There will also be an oceanic component, looking at the large water bodies that border North America. The program will include multi-scale, multi-data and multi-model integration with the common goal of characterizing the carbon cycle well enough to reduce the outstanding uncertainties in predictive models so that we will be better able to predict how biological sinks/sources might operate in the future.

An important component of the NACP will be the integration of ground based and aircraft measurements with remotely sensed data and biosphere-atmosphere models to extend results to the continent. There is a great deal of potential in using regional scale measurements and modeling to link the top down and bottom up approaches. Coupled regional atmospheric and land surface models preserve the property of spatial integration that is such an advantage in global atmospheric inversions while incorporating the explicit process based modeling at higher spatial resolutions that is necessary for predictive power. Further, regional scale observations from the NACP will permit the evaluation of the results of such coupled models. Once the strengths and weaknesses of these models have been assessed, they could be applied in other areas of the world where development of such a dense observational network is improbable. Thus a major challenge of the NACP will be to develop ways to integrate the disparate data streams in such a way that they will be of use to the modeling communities who are attempting to bridge the top down/bottom up divide; while a concurrent challenge to the modeling communities will be to determine the most appropriate ways to use the data.

Data assimilation has long been a staple of the atmospheric science community for weather forecasting and more recently for atmospheric inversion studies (Kalnay 2003). Data assimilation utilizes the concept of model-data fusion, in which observational data and models are used synergistically, to advance the predictive and explanatory power of both. Accordingly, the goal of data assimilation is the optimal combination of both process based models and observations.

Fundamental to data assimilation are the concepts of sensitivity and uncertainty. A common objective of data assimilation is the optimization of parameters and states in process models. As Kaminski et al. (Kaminski *et al.* 2002) have noted, this can only be achieved with a level of confidence when the parameter or state is sensitive to the observations. Further, quantifying the level of uncertainty in the models and observations is essential for robust results (Rodenbeck *et al.* 2003). Quantifiable uncertainty in both models and observations is used to constrain or relax the freedom of optimizations in realistic ways and to limit over-interpretation of results.

In the last few years, there have been many developments in the assimilation of CO_2 concentration data into atmospheric inversions including recommendations for reducing spurious interpretations through better data choices (Law *et al.* 2003a; Rodenbeck *et al.* 2003) and incorporating more highly time-resolved information (Law *et al.* 2003b). A relatively new development is that of Carbon Cycle Data Assimilation (CCDA; Kaminski *et al.* 2002; Kaminski *et al.* 2003), in which atmospheric concentration data and possibly other observational data is used to optimize a terrestrial biosphere model. In CCDA the goal is to improve the parameterization and predictive power of process models such that they provide more robust fluxes for atmospheric inversions, to quantify the uncertainty in the surface fluxes and models better, and to increase the amount and kinds of data that can be used to constrain atmospheric inversions (Kaminski *et al.* 2002). Initial work with a relatively simple terrestrial biosphere model (two parameters to optimize) has shown that both CO_2 concentration data and flux measurements (as from a tower) can provide powerful constraints on model parameters and the resulting predictions of fluxes (Kaminski *et al.* 2002). Work is in progress with a much more complex terrestrial biosphere model (Kaminski *et al.* 2003).

This dissertation addresses some of the issues likely to be faced in this new and challenging environment. In Chapter 2 a biophysical land surface model (SiB2.5; Baker *et al.* 2003) is introduced and its application to a temperate, forested region in the upper mid-west U.S. is discussed. SiB2.5 is a complex, non-linear, highly parameterized model that has been coupled to both a general circulation model and a regional atmospheric model. Recent work has suggested that when large numbers of parameters are required in models such as SiB2.5, there may be compensation among parameters which results in physically reasonable simulations across a wide range of parameter values "equifinality" (Franks *et al.* 1997; Franks 1998). Equifinality in land surface models precludes efficient optimization if indeed optimization is truly possible. In Chapter 3, the sensitivity of SiB2.5 to its parameterization is examined. Parameter sensitivity is assessed with respect to the ability of SiB2.5 to simulate fluxes of latent and sensible heat as well as the net ecosystem exchange of carbon as measured on a flux tower at both monthly and annual time scales. Parameters that do not contribute to the success of the model are identified and the remaining parameters and resultant distributions of those parameters are discussed. In Chapter 4, the results of the sensitivity analysis are used to assess the uncertainty in fluxes simulated with SiB2.5 attributable to the parameterization. Using a Bayesian framework, uncertainty bounds on the simulated fluxes are calculated at both the yearly and monthly timescale and the appropriateness of the time varying and time invariant parameterization of the model is discussed.

Models such as SiB2.5 often rely on remotely sensed data to represent surface processes and/or to provide state variables, providing the critical link between local, regional and global scales. The advantage of using remote sensing in such modeling efforts is its ability to provide parameter fields not easily measured, either temporally or spatially, at the ground surface. This advantage, however, can also be seen as a disadvantage in that the spatial and temporal resolutions of the data are fixed (30m, 250m, 1km, every 2 days, every 10 days, etc…) and may or may not be appropriate to represent the heterogeneity of the surface accurately. Further complicating matters, the effects of scale and subgrid scale heterogeneity are likely to be most important where small changes in parameterization lead to large changes in predictions. In Chapter 5 the uncertainty in regional simulations of latent and sensible heat flux and net ecosystem exchange resulting from differing representations of land surface heterogeneity is evaluated. A version of SiB2 coupled to a regional atmospheric model (RAMS; Cotton *et al.* 2003) is parameterized with two different representations of land surface heterogeneity for the upper mid-west United States. Simulations are conducted for three phenologically distinct time periods; late spring as soils thawed but before leaf out, the advent of leaf out and full summer vegetation. Uncertainty in regional flux simulations as well as local variability are discussed and compared to earlier results.

A central objective of contemporary carbon cycle science is to move from separate top-down and bottom-up approaches to more comprehensive spatial and temporal analyses that will both improve prediction and provide a strong foundation for policy formulation in the future. Thus, it is crucial to understand the sources of error, potential for bias and impacts of model and data choices as they provide greatly needed context. This dissertation addresses aspects of sensitivity and uncertainty in carbon cycle modeling that will be important for improving methodologies and interpretations of data, especially when programs such as the NACP come on-line.

CHAPTER 2

COMMON DATA AND MODEL

2.1 Introduction

Working across time and space scales requires the integration of many types of datasets. Presented in this chapter are descriptions of data and a model common to the experiments in this dissertation. These include a description of the site, meteorological forcing data and surface flux observations for the site, the principal land surface model and boundary conditions for that model. Together they provide the physical context for the experimental modeling work. Later chapters present additional data sets and models relevant to particular experiments.

2.2 Site Description

The WLEF-TV tower (90.28°W, 45.95°N) is located within the Chequamegon Nicolet National Forest in the northern highlands region of Wisconsin, approximately 100 Kilometers south of Lake Superior. Micrometeorological, eddy covariance and CO_2 concentration measurements have been made since 1995 on this tall tower (a 447 m tall TV tower; Figure 2.1.) (Bakwin *et al.* 1998; Berger *et al.* 2001; Davis *et al.* 2003). The vegetation of the region is composed of mixed northern hardwoods, upland Jack and Red pine, lowland conifers, aspen and wetlands (Figure 2.2) (WiDNR 1998; Burrows *et al.* 2002). Much of the area was logged from 1860-1920 and has since re-grown (USDAFS 2001). The topography is mild, with a local relief of up to 45 meters, and elevations reaching to approximately 450 meters above sea–level (Martin 1965; Burrows *et al.* 2002; Davis *et al.* 2003). The climate of the area is cool continental with a thirty-year (1971-2000) annual average mean temperature of 4.9 degrees C (-12.3 °C average in January; 19.6 °C average in July) and annual precipitation averaging 815 mm yr^{-1} (MRCC 2002). The area has mixed recreational usage year-round.

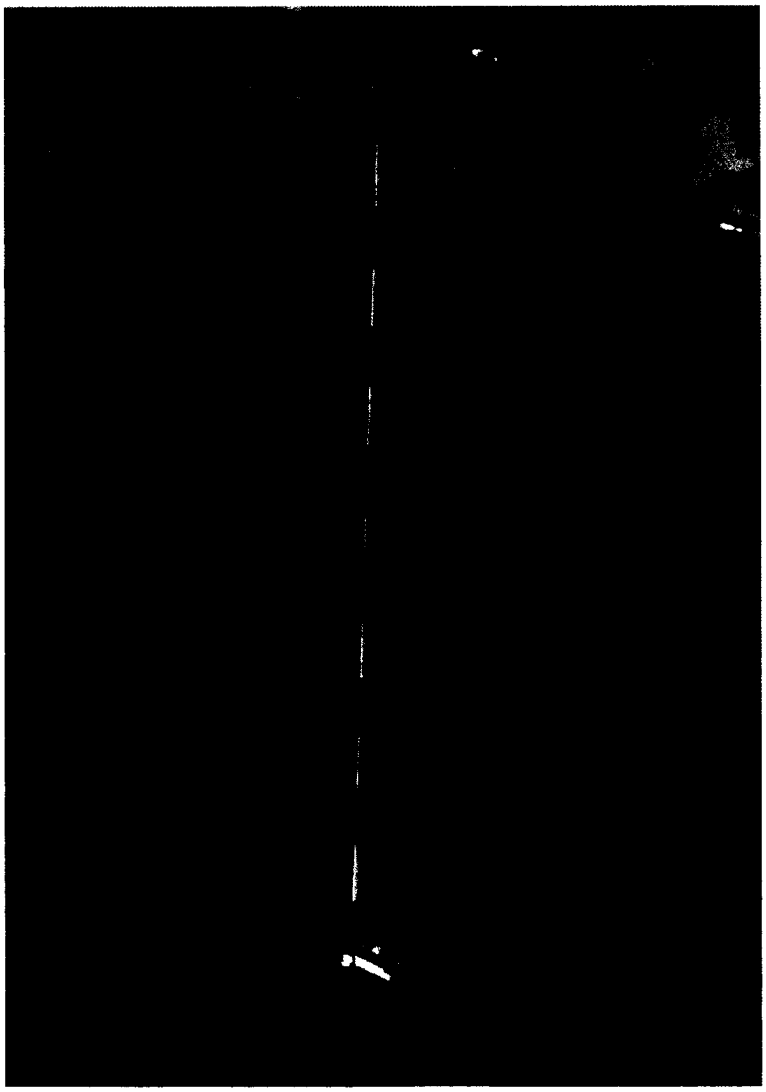


Figure 2.1. The WLEF tall tower and surrounding region.

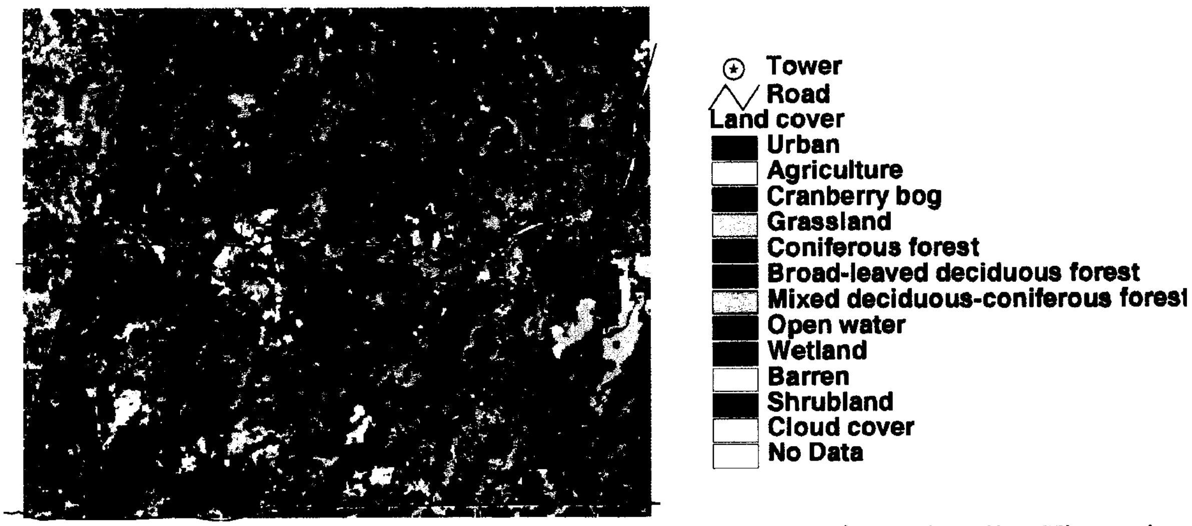


Figure 2.2. Land cover in the vicinity of the WLEF TV tower site Park Falls, Wisconsin (WiDNR 1998).

Studies are being undertaken at the tall tower and in the surrounding region to assess the exchange of CO_2 and energy between the forest and the atmosphere and the processes contributing to these fluxes (Davis *et al.* 2003). Additional studies address questions of scale (Burrows *et al.* 2002; Mackay *et al.* 2002), measurement of regional surface processes by remote sensing and the role of atmospheric boundary layer dynamics in regulating the carbon dioxide concentration near the ground (Yi *et al.* 2001; Denning *et al.* 2003). Collectively, these studies form the Chequamegon Ecosystem Atmosphere Study (ChEAS) (for details see http://cheas.psu.edu/)

2.3 Observational Data

2.3.1 *Meteorological data*

Micro-meteorological measurements have been made at the WLEF tall tower since 1995 (Bakwin *et al.* 1998; Berger *et al.* 2001; Davis *et al.* 2003). Wind speed and direction, air temperature and relative humidity are measured at 30m, 122m, and 396m. Atmospheric pressure, precipitation, photosynthetically active radiation (PAR) and net radiation are measured at the surface in the clearing surrounding the tower. For this study meteorological measurements for the year 1997 at 122m were used. These measurements were averaged at hourly intervals. Figure 2.3a,b shows the average monthly relative humidity at 122m and monthly maximum PAR. Incoming long-wave radiation was not measured, so the method of Idso (Idso 1981) was used to estimate long-wave radiation from air temperature and humidity.

A hierarchical sequence of methods was used to fill gaps in the meteorological data due to intermittent instrument or data logger failures (Baker *et al.* 2003). The filled meteorological variables used in this dissertation are: air temperature, dew point temperature and wind speed at 122m, atmospheric pressure, incoming long wave radiation, incoming short wave radiation and precipitation at the surface.

> "The gap filling sequence included, in order of preference: (1) temporal interpolation within the data stream for gaps of less than 2 hours, (2) use of data from an alternative height on the WLEF tower, with adjustment for systematic differences caused by height, (3) use of data from the nearby Willow Springs site, with adjustment for systematic differences caused by location, (4) use of an average for the time of day of missing measurements, using data from the previous and following 15 day period, and (5) use of an average for the time and date of missing measurements, using data from the previous and following 15 day period, in the same year and any other available year. In methods (2)-(5) the mean adjustment and averages were calculated, with standard deviations, using measurements at that time of day during the 15-day period before and after the time and date of the missing measurement. Methods (2)-(5) were not used if the variability of the adjustment or average was more than a threshold value (defined as standard deviation/mean = 2.0)." (Baker *et al.* 2003).

Figure 2.3d shows monthly average air temperature at 122 meters on the WLEF tower compared to the thirty-year mean at the Park Falls weather station (MRCC 2002).

Gaps in the precipitation data were filled directly with measurements from the Willow Springs site from May through November. During the wintertime (December through April), the WLEF and Willow Springs gauges had problems recording snowfall events, a problem typical of tipping bucket type gauges. To produce more accurate snowfall driver data, snowfall observations from nearby (13.3 km to 91.6 km) national weather service cooperative stations were used to determine wintertime precipitation. Substitute stations with available data were chosen in order of proximity to the tall tower. These daily observations were divided into equal hourly increments under the assumption that winter precipitation would generally be stratiform in nature (Figure 2.3 c).

2.3.2 *Flux Observations*

High frequency flux observations of latent (LE) and sensible heat (H) and net ecosystem exchange of CO_2 (NEE) have been made on the tall tower at three heights (30m, 122m and 396m) since 1995 (Bakwin *et al.* 1998; Berger *et al.* 2001; Davis *et al.* 2003). At each of these heights three-dimensional wind speed and direction, sonic temperature, relative humidity, and the mixing ratios of CO_2 and H_2O are sampled at 5hz and averaged over an hour. In addition, the mixing ratio of CO_2 and H_2O are measured at three other heights (11m, 76m and 244m). Corrections were applied to the flux data to account for storage effects and preferred values of NEE, LE and H were selected based on atmospheric conditions (Davis *et al.* 2003). The preferred values were selected to best represent the heterogeneous landscape in the area by maintaining a large, consistent flux footprint while lessening the influence of the cleared patch (radius 200m) underneath the tower.

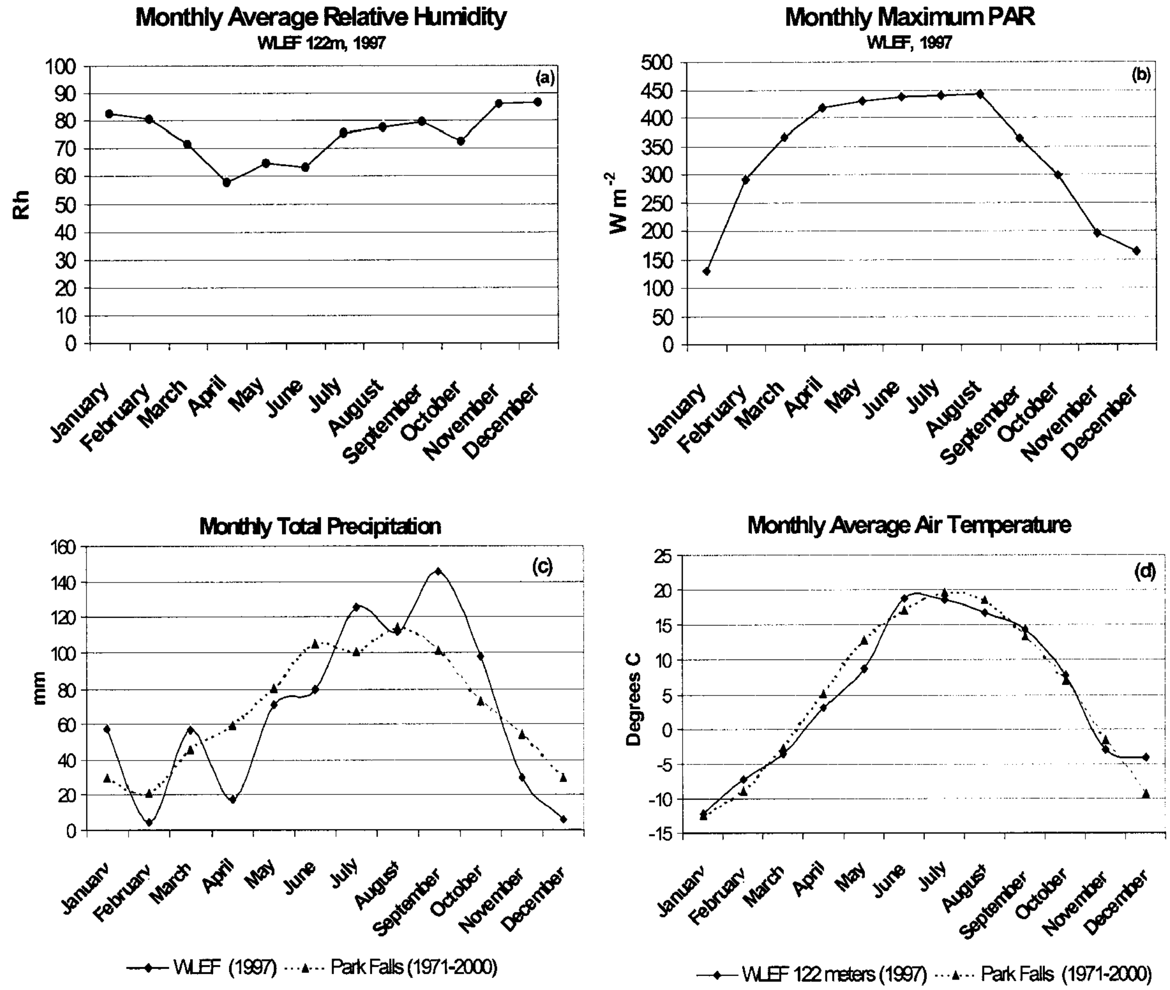


Figure 2.3. (a) Monthly average relative humidity at 122m, (b) monthly maximum PAR, (c) monthly total precipitation (d) and monthly average air temperature. Park Falls historical averages from the Midwestern Regional Climate Center (MRCC 2002).

Corrections for storage were computed by calculating the rate of change of the mixing ratio of a gaseous species with height from the 6 levels and adding the change of storage below a given measurement height to the turbulent flux of the species at that height. In the case of sensible heat, it is the vertically integrated enthalpy. The net ecosystem exchange of a gaseous species or air temperature is then:

$$F_0 = \int_0^{z_r} \frac{\partial \overline{s}}{\partial t} dz + \overline{w's'}_{z_r} = F_{Sst} + F_{Stb} \qquad (2.1)$$

where s is a mixing ratio of a gaseous species ($[CO_2]$,$[H_2O]$) or air temperature (T), z is the height, z_r is the measurement height, w is the vertical velocity, F is the net flux, *st* indicates storage and *tb* turbulent flux. Perturbations are indicated by primes and the time average mean by the overbar. Additionally, for LE the covariance of [H_2O] is multiplied by λ, where λ is the latent heat of vaporization and for H the covariance of T is multiplied by ρC_p where ρ is the density of air and C_p is the specific heat of air at constant pressure.

Preferred values of NEE, LE and H were then chosen based on stability. When surface fluxes of sensible heat were greater than 100 Wm^{-2}, conditions were considered to be strongly unstable and measurements were taken from 122m, 396m or an average of the two. When surface fluxes of sensible heat were less than 100 Wm^{-2}, conditions were considered to be slightly unstable to stable and measurements were taken from 30m. If data were missing from the favored level, data from one or two of the other levels were used. Short gaps of missing data, one to two hours, were interpolated. A dataset where larger gaps were filled was not used in this dissertation. See Davis et al.. (Davis *et al.* 2003) for a complete description of the methodologies used for both storage and stability corrections.

For the work presented in this dissertation, flux data from the year 1997 were used. 1997 was chosen because conditions were close to the thirty-year average for temperature and precipitation and because it was the first full year of flux data collected at WLEF. The data were also thoroughly analyzed in a paper by Davis et al. (Davis *et al.* 2003). Approximately twenty-nine percent of the observations of NEE and thirty-three percent of latent and sensible heat flux observations for May were missing. Approximately thirty-seven percent of the observations of NEE and forty percent of latent and sensible heat flux observations for August were missing. Overall, 90% of NEE observations and 85% of the possible observations of sensible heat flux and latent heat flux were available. Figure 2.4 shows a monthly mean diurnal plot of observed latent and sensible heat flux and NEE plotted with simulations for July 1997. See appendix one for monthly mean diurnal plots for all 12 months.

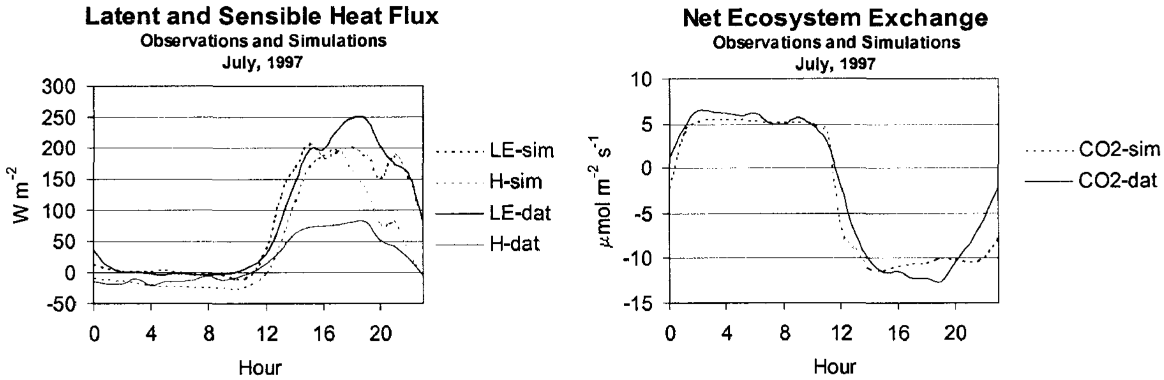


Figure 2.4. Monthly mean diurnal cycles of latent and sensible heat flux and net ecosystem exchange, simulations (-sim) and observations (-dat), for WLEF, July 1997 in GMT.

2.4 Model Description

The Simple Biosphere Model version 2 (SiB2; Sellers *et al.* 1996b) is a single canopy-layer scheme describing the transfers of heat, water and carbon in the soil-vegetation-atmosphere continuum. SiB2's physical model of radiation balance, heat and water transport conserves mass and energy. SiB2 incorporates leaf-level physiology controlling photosynthesis (Farquhar *et al.* 1980; Collatz *et al.* 1991; Collatz *et al.* 1992) and the Ball-Berry (Ball *et al.* 1987) description of stomatal behavior which links water loss and carbon assimilation. Canopy scale calculations of carbon and water fluxes are estimated from leaf level physiology and exchange using a canopy integration factor (Π) that is related to the extinction of photosynthetically active radiation (PAR) through the canopy (Sellers *et al.* 1992a). Π reflects the assumption of continuously varying stomatal conductance, carbon assimilation, and rubisco velocity (Vmax) vertically in the canopy in response to the time-mean distribution of light, with nitrogen allocation responding accordingly (Field and Mooney 1986). The integration factor allows for a direct relationship between canopy scale fluxes and larger spatial scales because the fraction of absorbed photosynthetically active radiation (f_{PAR}) by a canopy is near linearly related to the simple ratio (SR) (Sellers 1985; Sellers 1987):

$$SR = \frac{IR}{R} = \frac{(1+NDVI)}{(1-NDVI)} \tag{2.2}$$

$$NDVI = \frac{IR-R}{IR+R} \tag{2.3}$$

where IR is near infrared reflectance and R is red reflectance. Near infrared and red reflectance, and thus the simple ratio and the normalized difference vegetation index (NDVI; see eq. 2.3), are commonly observed by remote means.

In SiB2 the annual sum of the respiratory loss of carbon from the soil is assumed to be equal to the annual sum of the canopy net carbon assimilation. Respiratory losses are parameterized diurnally and seasonally by partitioning respiration between the six soil layers and the surface and regulating the respiration rate through temperature, moisture and soil texture (Raich and Nadelhoffer 1989; Denning *et al.* 1996a). Therefore, when run at an annual time scale with this parameterization of soil respiration there is a net zero balance of carbon into and out of the system. At shorter time scales, however, carbon is not conserved. At WLEF in 1997 a net zero balance of carbon into and out of the system is a reasonable assumption, the NEE of CO_2 was 16 ± 19g C m^{-2} yr^{-1} (Davis *et al.* 2003).

The use of a single canopy-layer scheme, and in particular the methods used to scale processes from leaf to canopy, permit SiB2 to be driven by remotely sensed estimates of vegetation properties (Sellers *et al.* 1996a). This in turn simplifies the parameterization of the model while better representing the variability of vegetation in space than models that assign uniform properties to all grid cells of a particular vegetation class. It also allows SiB2 to be run at a wide range of spatial scales. In SiB2, primary parameters affecting land surface-atmosphere exchanges include canopy and soil structural and optical properties and physiological parameters controlling photosynthesis and respiration. The most recent version, SiB2.5 (Baker *et al.* 2003), includes prognostic equations describing transfers of heat, water and carbon that also account for canopy air-space storage. Also incorporated into SiB2.5 was a six-layer soil temperature model, an update from the three-layer model in previous versions. SiB2 and SiB2.5 can be run as single point simulations as well as coupled to regional and global atmospheric models (Denning *et al.* 1996a; Denning *et al.* 1996b; Randall *et al.* 1996; Colello *et al.* 1998; Baker *et al.* 2003; Denning *et al.* 2003; Nicholls *et al.* 2004).

2.4.1 *Model boundary conditions*

Boundary conditions for SiB2 and SiB2.5 characterize land surface conditions at a location using a combination of land cover type, monthly maximum Normalized Difference Vegetation Index (NDVI) and soil properties. Two types of parameters represent the boundary conditions, time invariant biophysical parameters (40) (Table 2.1a) and time-varying biophysical parameters (8) (Table 2.1b). Time invariant biophysical parameters include quantities such as canopy height, leaf angle distribution, leaf optical transmittance and the physiological parameters related to photosynthesis and respiration, as well as soil hydraulic and thermal properties (see Table 2.1a for a full list). The majority of these parameters are defined through literature values and are assigned by biome, of which there are 9 distinct types in SiB2 and SiB2.5 (see Sellers *et al.* 1996a). In the standard version of SiB2 the time invariant soil hydraulic and thermal properties are assigned by soil textural class. However, in this study these are calculated from the percent sand and percent clay of the soil using the equations of Cosby et al (Cosby *et al.* 1984) as modified by Bonan (Bonan 1996). This modification permitted better representation of the spatial variability in soil physical characteristics. The soil parameters and their equations are listed in Table 2.2. Time invariant parameters which are used in calculations of heterotrophic respiration from the soil are estimated from the percent clay of the soil using curves fit to data in Raich et al (Raich *et al.* 1991) (K. Schaeffer, Colorado State University, personal communication). These parameters and their equations are listed in Table 2.3.

Time-varying biophysical parameters include quantities such as leaf area index, fractional absorbed photosynthetically active radiation and surface roughness length (see Table 2.1b for a complete list). Most of these parameters are estimated either directly or indirectly from NDVI data using the methods of Sellers and Los (Sellers *et al.* 1996a; Sellers *et al.* 1996b; Los *et al.* 2000). GMUDMU, the time mean leaf projection, is not estimated from NDVI but varies at a monthly time step depending on the solar declination and latitude.

Table 2.1a. Standard time invariant boundary conditions for the WLEF site

Parameter	***Definition***	***WLEF***
Time invariant parameters		
Z2	Canopy top height (m)	20.0
Z1	Canopy base height (m)	10.0
ZC	Canopy inflection height (m)	15.0
VCOVER	Vegetation cover fraction	0.9875
CHIL	Leaf angle distribution factor	0.125
ROOTD	Rooting depth (m)	1.5
PH	½ critical leaf water potential (m)	-200.0
TRAN11	Green leaf transmittance (PAR)	0.05
TRAN21	Green leaf transmittance (NIR)	0.15
TRAN12	Brown leaf transmittance (PAR)	0.001
TRAN22	Brown leaf transmittance (NIR)	0.001
REF11	Green leaf reflectance (PAR)	0.07
REF21	Green leaf reflectance (NIR)	0.38
REF12	Brown leaf reflectance (PAR)	0.16
REF22	Brown leaf reflectance (NIR)	0.42
VMAX0	Rubisco velocity of sun leaf (mol m^{-2} s^{-1})	7.5E-5
EFFCON	Quantum efficiency (mol mol^{-1})	0.08
GRADM	Conductance-photosynthesis slope parameter	9.0
BINTER	Minimum stomatal conductance (mol m^{-2} s^{-1})	0.01
ATHETA	Light and rubisco coupling parameter	0.98
BTHETA	Light, rubisco and CHO sink parameter	0.95
TRDA	Respiration temperature response (K^{-1})	1.3
TRDM	Respiration inhibition ½-point temperature (K)	328.16
TROP	Respiration optimum temperature (K)	298.16
RESPCP	Leaf respiration fraction of Vmax	0.015
SLTI	Photosynthesis low temperature response (K^{-1})	0.2
SHTI	Photosynthesis high temperature response (K^{-1})	0.3
HLTI	Photosynthesis low temperature inhibition ½-point (K)	280.66
HHTI	Photosynthesis high temperature inhibition ½-point (K)	307.16
BEE	Slope of the retention curve	5.29
PHSAT	Soil water potential at saturation (m)	-0.25
SATCO	Hydraulic conductivity at saturation (m)	3.4E-6
POROS	Water content at saturation (porosity)	0.44
SLOPE	Cosine of mean terrain slope	0.1736
WOPT	Optimal percent of soil saturation for respiration	62.12
ZM	Skewness exponent of respiration vs. soil water	0.362
WSAT	Coefficient for respiration rate at soil water saturation	0.537
SODEP	Soil depth (m)	2.0
SOREF1	Soil reflectance (PAR)	0.11
SOREF2	Soil Reflectance (NIR)	0.22

Table 2.1b. Standard time varying boundary conditions for the WLEF site, 1997.

Time varying parameters January-December, 1997													
		Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
APAR	Fraction of absorbed photosynthetically active radiation	0.42	0.3	0.2	0.47	0.69	0.94	0.95	0.95	0.95	0.89	0.63	0.51
ZLT	Leaf area index	0.58	0.39	0.29	0.78	1.53	4.04	7.42	6.7	4.97	2.77	1.12	0.71
GREEN	Greenness fraction	0.65	0.65	0.54	0.9	0.95	0.98	0.99	0.88	0.78	0.54	0.53	0.89
Z0D	Surface roughness length	0.5	0.39	0.32	0.59	0.81	1.02	1.02	1.03	1.03	0.96	0.71	0.56
ZP_DISP	Zero plane displacement height	15.3	15.1	14.9	15.4	15.7	16.2	16.6	16.6	16.4	16.0	15.6	15.4
CC1	Boundary layer resistance coefficient	41.6	58.3	78.6	32.3	18.7	9.2	6.4	6.8	8.1	11.9	24.0	35.1
CC2	Canopy layer resistance coefficient	281.0	269.6	261.6	292.2	331.2	462.8	665.9	619.5	515.8	395.0	310.3	288.2
GMUDMU	Time mean leaf projection	1.10	1.03	0.94	0.88	0.84	0.83	0.84	0.87	0.92	1.01	1.09	1.13

Table 2.2. Equations for time invariant soil hydraulic and thermal parameters.

Name	*Definition*	*Equation (number)*	
PHSAT	Soil water potential at saturation (m)	$=\dfrac{-10\times10^{(1.88-0.0131\times Fsand)}}{1000}$	(2.4)
POROS	Water content at saturation (porosity)	$=0.489-0.00126\times Fsand$	(2.5)
SATCO	Hydraulic conductivity at saturation (m)	$=\dfrac{0.0070556\times10^{(-0.884+0.0153\times Fsand)}}{1000}$	(2.6)
BEE	Slope of the retention curve	$=2.91+0.159\times Fclay$	(2.7)

* *Fclay* and *Fsand* are the fraction of clay and sand in the soil as a whole number.

Table 2.3. Equations for time invariant heterotrophic respiration parameters.

Name	*Definition*	*Equation (number)*	
WOPT	Optimal percent of soil saturation for respiration	$=(-0.08\times fclay^2+0.22\times fclay+0.59)\times100$	(2.8)
ZM	Skewness exponent of respiration vs. soil water	$=-2\times fclay^3-0.4491\times fclay^2+0.2101\times fclay+0.3478$	(2.9)
WSAT	Coefficient for respiration rate at soil water saturation	$=0.25\times fclay+0.5$	(2.10)

* *fclay* is the fraction of clay in the soil in percent.

Listed in Table 2.1a,b are the standard boundary conditions for 1997 for the WLEF site. Boundary conditions for the single point model characterize the site for a typical flux footprint of tower measurements at 122m. At 122m, a region of at least 24 km^2 would incorporate the primary source region of the measured fluxes, given the average wind speed (6 ms^{-1}) at the site. This dimension was based on simulations of the tower flux footprint using lagrangian particle dispersion techniques conducted by Marek Uliasz (Colorado State University, personal communication). For WLEF, the vegetation was classified as mixed deciduous and coniferous forest (SiB2 class 3), and the soil as consisting of 37% percent sand and 15% percent clay, which was determined from the STATSGO soil database (Staff 1994). Time varying boundary conditions are calculated from an annual series of monthly mean normalized difference vegetation index measurements (Los *et al.* 2000), in this case calculated from 8 km Advanced Very High Resolution Radiometer (AVHRR) data for 1997.

Estimated SiB2.5 initial conditions for the WLEF site on January 1, 1997 are listed in Table 2.4. Initial conditions characterize the state of the land surface at the beginning of a given simulation and include conditions such as soil temperature and moisture and snow and/or water present on the surface. To create a reasonable soil moisture profile and deep soil temperatures, we estimated these initial condition values through a 12 year 'spin-up' with SiB2.5, cycling through three years of meteorological forcing data (1995, 1996 and 1997). Data in the table represent the condition of the model at midnight on December 31.

Table 2.4 Standard initial conditions for the WLEF site, 1/1/1997.

Initial Condition	*Value*
Canopy Temperature	257.6 K
Ground Temperature	258.5 K
Deep Soil Temperature – Level 1 (lower)	276.8 K
Deep Soil Temperature – Level 2	274.2 K
Deep Soil Temperature – Level 3	273.4 K
Deep Soil Temperature – Level 4	265.6 K
Deep Soil Temperature – Level 5	262.0 K
Deep Soil Temperature – Level 6 (upper)	259.9 K
Water on vegetation surface	0.0 m
Water on ground surface	0.0 m
Snow on vegetation surface	0.0 m
Snow on ground surface	0.02 m
Soil water percent of saturation level 1	0.71
Soil water percent of saturation level 2	0.71
Soil water percent of saturation level 3	0.73
Stomatal resistance	182572 mol m^{-2} s^{-1}
Turbulent kinetic energy	4.17

2.5 Model Performance

Flux output from SiB2.5 was averaged from the 10-minute model time step to hourly output, to correspond with the observational time step. Monthly mean diurnal plots of latent and sensible heat flux and NEE, both simulations and observations for all twelve months, can be found in Appendix A. These plots represent the capability of the model to capture the salient features of the diurnal cycle of fluxes at the monthly time scale using standard model parameter values.

The proficiency with which SiB2.5 simulates fluxes of sensible and latent heat and NEE at WLEF varies seasonally, monthly and diurnally. In 1997 at WLEF soils thawed in mid-April, on the 14th, and didn't re-freeze again until November. Leaf out followed the thaw, occurring around May 30th and leaf fall preceded the freeze, occurring around October 10th (Davis *et al.* 2003). The monthly mean diurnal plots show that sensible heat flux is simulated quite well during those months when leaves are not present on the canopy from October through May, but the model quickly begins to overestimate sensible heat flux once leaves are on the canopy in June. The over-estimation persists through the summer until October, when the leaves fall. For simulations of latent heat flux, the monthly mean diurnal plots show greater agreement with observations in January and in May through December, than they do in late winter/early spring (February through April).

An important consideration to note is that the energy balance estimated from flux towers rarely closes. A recent analysis of FLUXNET data spanning 22 flux tower sites and 50 site years estimates that the discrepancy between available energy and turbulent energy fluxes was 20% on average (Wilson *et al.* 2002). SiB2 and SiB2.5 conserve mass and energy so it is likely that at least some of the discrepancy between observed and simulated fluxes is due to the lack of closure in the observations, though probably not all (Baker *et al.* 2003). Calculations of the Bowen ratio from the monthly mean diurnal calculations of LE and H for observations and simulations indicate that May is the only month where the observations and simulations agree well. In the winter months, the *observations* have a higher Bowen ratio than the simulations. In the summer months the *simulations* have a higher Bowen ratio than the observations. Thus the imbalance between the simulations and observations is not consistent with a systematic measurement problem. Baker et al. (Baker *et al.* 2003) suggest the lack of representation of wetlands in SiB2.5, which constitute approximately 30% of the landscape around WLEF (Mackay *et al.* 2002), may contribute to the overestimation of sensible heat flux in the summer months. However, the latent heat is not underestimated. It is possible the overestimation of latent heat flux in the late winter/early spring by the model is due to poor representation of snow cover in the model. In SiB2.5, once snowpack is greater than 6mm water equivalent, the entire surface is considered covered with a layer of snow. If the snow was patchy in the actual footprint, the result could be an overestimation of LE from evaporation and sublimation from the modeled surface as compared to the observations.

The monthly mean diurnal plots indicate that simulations of NEE are quite reasonable during most times of the year, with the exception of May and September. In May, the model overestimates NEE for daytime fluxes and in September it underestimates NEE for daytime fluxes. One possible explanation for this is that in SiB2.5, leaf area is calculated for the mid point of each month based on the maximum monthly NDVI. This mid-month value of NDVI is then interpolated between months to give daily variation in leaf area. Leaf out occurred at the very end of May so the whole month of May should have had very low leaf area. However, since leaf area from the midpoint of May would have been interpolated from a low value to a high value (the mid-point of June), there would have been an overestimation of leaf area for the second half of the month. Similarly, since leaf fall was in the beginning of October, leaf area for the second half of September would have been interpolated from a time of high to low leaf area perhaps causing an underestimation of NEE. Figure 2.5 shows the time series of NDVI used to calculate the time varying parameters for the point simulation of WLEF. A large jump in NDVI is visible between May and June and again between September and October, which is reflected in the monthly values of absorbed photosynthetically active radiation and leaf area index reported in Table 2.1b. The midlevel values of NDVI seen in April, May and October most likely represent the presence of coniferous vegetation. That the values of NDVI drop in the winter months, despite the presence of coniferous vegetation, is probably due to the presence of snow on the ground, which has the effect of reducing NDVI because of it's bright reflectance in visible wavelengths.

The root mean square difference (RMSD) between observed and simulated fluxes for all twelve months, where:

$$RMSD = \sqrt{\frac{\sum (obs - est)^2}{(n-1)}} \tag{2.11}$$

is presented in Table 2.4. The RMSD errors are calculated on a time step basis, so, unlike the monthly mean diurnal plots that show the bias between observations and simulations on a monthly time scale, the RMSD error reflects the average hourly error between simulations and observations. The RMSD does not provide information on bias because the direction of the error is not given. It does, however, give an indication of the variability of the observations and the model's ability to capture that variability, which is smoothed out in the monthly mean diurnal composites. The RMSD errors follow a somewhat different pattern than those of the monthly mean diurnal plots. The RMSD error in sensible heat begins to increase earlier in the season, at soil thaw but before leaf out, and decreases at leaf fall. There are also fairly large errors in January. The RMSD errors in latent heat show larger errors May through September, when the leaves are on, and the smallest errors in the winter months. The RMSD errors for NEE begin to increase in April, reach their peak in June and decrease steadily beginning in August, resulting in the highest RMSD errors in June, July and August. Larger errors in the summer months are likely due to more variability in the measurements caused by surface turbulence.

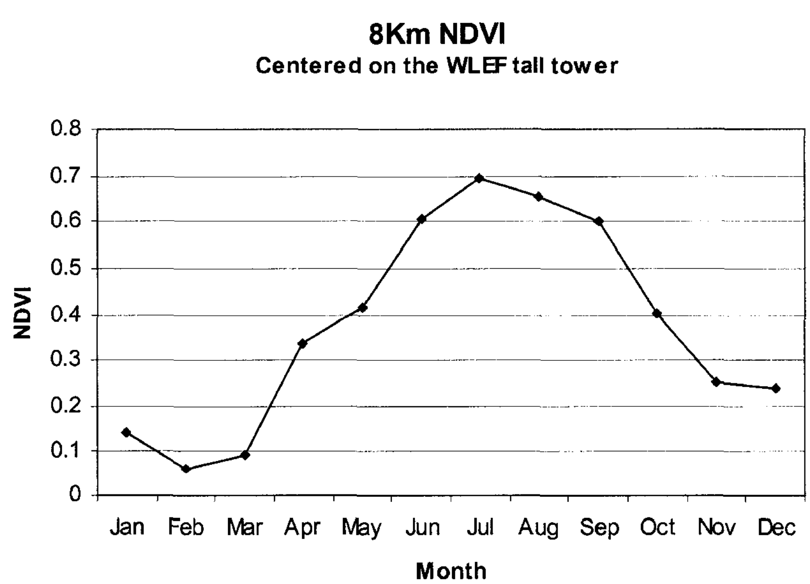


Figure 2.5. Time series of 8km NDVI for 1997 for the WLEF site. NDVI was calculated from AVHRR by S. Los using the FASIR methodology (Los *et al.* 2000).

Table 2.4. Root mean square difference (RMSD) between observed and simulated fluxes for 1997 for each month and for all months combined. Results are shown in units of flux, $\mu mol\ m^{-2}\ s^{-1}$ for NEE and $W\ m^{-2}$ for LE and H.

Month	*NEE*	*LE*	*H*
January	0.90	11.57	49.74
February	1.10	21.48	33.28
March	0.98	29.80	33.54
April	1.93	38.86	47.21
May	3.63	44.14	48.15
June	5.18	66.10	44.39
July	4.89	83.91	59.18
August	4.94	80.82	62.93
September	3.37	64.02	63.55
October	2.23	43.42	40.58
November	1.47	16.42	30.61
December	0.93	11.80	20.38
All Months	3.02	48.05	45.37

2.6 Discussion

This chapter presents the physical context for much of the experimental modeling work in this dissertation. Descriptions of the site, meteorological forcing data and surface flux observations for the site, the principal land surface model and boundary conditions for that model were provided as well as an evaluation of the performance of the model at the WLEF tall tower site in Park Falls, Wisconsin. It is clear that even given our best effort at observing surface fluxes, formulating models to predict those fluxes and characterizing the land surface there are inherent errors and biases. If we are to understand and gauge the predictive power of the models we use, it is necessary to examine the origins of this uncertainty. The chapters that follow address temporal and spatial uncertainty in modeling the land surface, building on the framework presented here.

CHAPTER 3

SENSITIVITY ANALYSIS

3.1 Introduction

Simulation models of ecosystem processes are powerful tools for gaining insights about the function and dynamics of ecosystems, for hypothesis testing and for prediction. With simulation models, one can test hypotheses that would be impractical to test or even improbable in the real world. Simulation models can also be used to predict what could happen in the future or why things happened in the past based on current theories and mechanistic understanding of biological and ecological processes.

However, the process of modeling is imperfect. Models can never fully represent every component and process in a complex ecosystem (it would be impractical and unproductive to do so), meaning that approximations and compromise solutions are always necessary at some level. It is a conundrum, to test hypotheses regarding ecosystem process with a model that is an imperfect representation of complex physical and physiological relationships. Yet in the end it is the iteration between model formulation and testing against observations that moves science forward. Careful revisiting of model structure, process representation and observations when expectations are not met contributes to the strength of modeling. Eventually the most critical aspects of a system present themselves for further investigation.

An important characteristic of simulation models is that they utilize parameters. Parameters in models are either measurable terms that are used to define the system and determine its behavior or are terms that represent important processes which for some reason (practical or theoretical) cannot be described in detail. Like models, parameters themselves are imperfect. If a parameter value is poorly defined, if it cannot be accurately measured or if it is not sufficiently representative it will introduce uncertainty into modeled results depending on the sensitivity of the model to that parameter. Complex models with many parameters (e.g. >20) can also be over-parameterized. That is, they contain parameters that contribute little to the simulation but have a cost associated with them in terms of computation time or definition. A consequence of over parameterization in complex models is a tendency towards equifinality. A model exhibits equifinality when compensation between parameters can result in equally good simulations across a wide range of parameter sets (Fedra *et al.* 1981; Spear *et al.* 1994; Franks *et al.* 1997; Franks 1998; Beven and Freer 2001). A tendency towards equifinality leads to increased uncertainty in predictions as a result of uncertainty in parameterization.

Biophysical land surface models (LSMs), which simulate fluxes of energy and mass between the atmosphere and the land surface, are used to study land surface dynamics, to generate surface conditions in atmospheric models and to elucidate the processes contributing to land surface fluxes. Typically these models are quite complex and can easily have more than 40 input parameters, putting them at risk of over-parameterization and equifinality. At present, biophysical LSMs coupled to regional and global atmospheric models are playing a large role in advancing our understanding of carbon cycle dynamics and associated changes in weather and climate (Denning *et al.*

1995; Pielke *et al.* 1998; Collatz *et al.* 2000; Cox *et al.* 2000; Bounoua *et al.* 2002). Because of their importance in determining potential trajectories for global and regional climate change, it is important that we understand the strengths and weaknesses of these models and how they could be improved.

An effective approach for fully exploring a model and its parameter space is the sensitivity analysis. Analyzing the sensitivity of model simulations and predictions to model parameterization can lead to insights about model structure and system behavior that would not ordinarily be evident simply from simulation results (O'Neill *et al.* 1982; Franks *et al.* 1997). Unexpected significance of a parameter or lack of significance can lead to questioning of our mechanistic understanding and the relative role of different processes. Sensitivity analyses are also useful for guiding research priorities. Knowing which parameters are most influential in a model can help focus research efforts in those areas where more attention would bring proportionately more benefit in terms of predictive accuracy (Hornberger and Spear 1981).

This chapter presents a sensitivity analysis of the biophysical land surface model, SiB2.5. A methodology is presented that examines the ability of SiB2.5 to simulate fluxes of latent and sensible heat and the net ecosystem CO_2 exchange for the WLEF tall-tower site in Park Falls, Wisconsin (described in detail in chapter 2). Utilizing this methodology we consider the sensitivities of SiB2.5 to parameters that are evaluated using remotely sensed data, literature values and measurements. The model sensitivities are analyzed through time, on a monthly basis and for an entire year of simulations. Several questions regarding the parameterization of SiB2.5 specifically and of models in general are addressed. They are:

1. Are there parameters that can be readily identified through sensitivity analysis that, if more accurately known, would improve the results of the model?
2. Does model sensitivity to parameterization vary in time? If so what might tell us about the system we are modeling?
3. Can non-influential parameters be fixed without detriment to model simulations?
4. Does the simulation of NEE and comparison to flux measurements help to constrain model parameterization and thus the simulation of energy and water fluxes as well?

3.2 Background

Simple models of land surface heat fluxes that were once thought adequate for large-scale weather and climate modeling have grown more complex as it has been recognized that more realistic representation of land surface hydrology and biophysics improves predictions of atmospheric phenomena (Sellers *et al.* 1997a). The next generation of models, more focused on long-term predictive power, includes processes such as competition among plant functional types, the changing biogeography of vegetation and the biogeochemistry of land surface-atmosphere exchanges. As more and more phenomena are represented, the structure of models becomes more complicated, parameterizations more complex and models more vulnerable to equifinality. However, it is also possible that adding more physical realism to models will help to constrain model prediction closer to reality because of better representation of the interactions and feedbacks that in nature constrain ecosystem functions (Xia *et al.* 2002). An effective way to evaluate the proficiencies and inefficiencies of a model is to use sensitivity analyses to measure the response of simulated model outputs to changes in parameterization, initial conditions, environmental driver data and model structure.

A variety of methodologies exist to explore the various sensitivities of a model and there is considerable debate over the best approach (Saltelli 1999). The choice of methodologies depends on the balance between the desired outcome, the computational expense and the limitations of the method. A common approach, especially with LSMs, is what Saltelli (Saltelli 1999) terms the "one at a time" (or OAT). In this approach, a baseline scenario of a model is chosen as well as parameters of interest. Each parameter is then varied independently by a specified deviation from its base value (typically by ten or twenty percent) and the resulting model outputs are analyzed for differences (e.g. Potter *et al.* 2001; El Maayar *et al.* 2002). Another common approach is to use partial derivatives where the effect of x on y is calculated for one or many parameters in a forward model (e.g. Dale *et al.* 1988). Both of these approaches assume linearity, in the model or computed outputs, which is rarely the case in LSMs. Thus, the interpretation of results is limited because non-linear interactions between parameters are not considered and claims that one parameter is more sensitive than another cannot be sustained (Saltelli 1999).

Factorial methods (Henderson-Sellers 1993) address some of the problems mentioned above. In a full factorial experiment all combinations of factors/parameters are used to create an interaction matrix. Some number of perturbation levels is chosen and the model is run for these perturbation levels for all combinations of factors. In this way, parameter interactions are directly addressed. However, factorial methods are very expensive when a large number of parameters are involved. With two perturbation levels and a full factorial experiment, the number of runs necessary would be 2^n, where n is the number of factors and 2 represents the two perturbation levels. Further, a limited number of scenarios are possible when the parameters are limited to only high and low values, yet adding perturbation levels greatly increases the cost and complexity of the method. One way to get around the number of necessary runs is to use a fractional factorial design, where only a half or a quarter of the simulations of the full factorial are run, but this creates other issues regarding the representativeness of the chosen runs and the loss of higher order interactions. Methodologies such as the factorial, OAT and the forward partial derivatives, where specified steps determine parameter ranges, are very sensitive to extreme parameter values because they don't consider the possibility that different parameters or factors might be influential over different ranges. The result is that such techniques look at the influence of specific parameters on the outcome of a particular simulation rather than at the true sensitivity of the model.

Monte Carlo methodologies allow a more thorough assessment of non-linear interactions between parameters (Fedra *et al.* 1981; Hornberger and Spear 1981; O'Neill *et al.* 1982; Franks *et al.* 1997; Franks 1998; Bastidas *et al.* 1999; Beven and Freer 2001). In typical Monte Carlo-style analyses, a large number of model simulations are performed on randomly generated parameter sets and the resulting model outputs are compared with observations to generate a measure of goodness of fit. This goodness of fit measure is then used in a variety of ways to analyze the sensitivity of the model. The parameter sets are drawn from distributions of the parameter that are defined by the modeler. The distributions can be uniform if there is no a priori knowledge of the distribution of the parameter or can take a form such as a normal distribution if there is information on the mean and variance of a parameter. Realistic parameter ranges prevent extreme simulations while allowing a large parameter space to be explored. In this way, the non-linear interactions between parameters become evident in the computed outputs observed in the many simulations. A further advantage is that these types of analyses can rather easily be adapted for uncertainty analysis and calibration. However, Monte Carlo style analyses do have drawbacks. They can be computationally expensive, particularly for models that contain a large number of parameters, since they can require thousands of model runs. Further, it can be difficult to obtain an exact measure of variance attributable to a single parameter when many parameters are being explored, resulting in a more structural interpretation of model sensitivity than a quantitative assessment of parameter sensitivity (O'Neill *et al.* 1982).

An additional class of sensitivity analyses explicitly accounts for the contribution to the total variance of the simulation by each parameter simultaneously. Two techniques that fall into this category are the Fourier Amplitude Sensitivity Test (FAST) (Collins and Avissar 1994) and adjoint sensitivity analysis. In the FAST technique, as in Monte Carlo methodologies, sets of input parameters are drawn from predetermined distributions of parameters. Rather than randomly drawing from these sets, in FAST each parameter is assigned a frequency that is used to assemble the different parameter sets. Fourier analysis is then used on the model output to separate model response from parameter response using the frequencies assigned to the individual parameters. The contribution of a particular parameter to model output variance is then determined. The FAST technique requires fewer model runs than Monte Carlo techniques, for example only 411 simulations for 10 parameters (Collins and Avissar 1994), but requires that the parameters to be analyzed are independent of each other, which is never the case in LSMs. Further, interactions and covariance between parameters are not explicitly handled.

Adjoint sensitivity analysis (Hall *et al.* 1982; Errico and Vukicevic 1992; Margulis and Entekhabi 2001b; Margulis and Entekhabi 2001a) uses differential operators derived from the equations that constitute a model to calculate the sensitivity of a scalar quantity, e.g. a model output or model error with respect to observations, to model parameters. Model parameters are perturbed and sensitivity is measured by the contribution of each parameter to the total variance of the chosen scalar. Adjoint models can be a very elegant approach to sensitivity analysis, however there are some limitations. Adjoint models are in general very sensitive to the choice of scalar (Errico and Vukicevic 1992). In adjoints of complex LSMs this is made worse because many parameters covary and interact with each other, making solutions non-unique. Also, the results of the adjoint model become unreliable if parameters are perturbed too far from their base state, violating the assumption of linearity in the adjoint, which might limit the relevance of the results (Gardner *et al.* 1981). Further, the highly non-linear nature and the lack of mathematical tractability of complex land surface models can make the formulation of an adjoint prohibitive.

The strengths and flexibility of Monte Carlo style analyses are well suited to the problem of analyzing the sensitivity SiB2.5 to its parameterization. SiB2.5 is a complex, highly non-linear model with a large number of input parameters (46) and therefore interactions and compensation between parameters are expected to be important. Some of these parameters are not decisively known and so a wide range of possible (but realistic) values for them should be tested. In Monte Carlo style analyses it is straightforward to evaluate whether a parameter has been specified badly by considering the posterior distribution of the parameter as it relates to good simulations, using other methods it is not. Finally, even though Monte Carlo methodologies are computationally expensive, they are relatively simple to set up and easy to run in a distributed computing system.

3.3 Methods

3.3.1 *Parameter set generation*

For this experiment forty-six SiB2.5 parameters were randomly varied within physically reasonable ranges (personal communication J. Collatz, N. Hanan; Sellers *et al.* 1996a), assuming a uniform distribution for all parameters (Table 3.1). The full suite of 40 time invariant parameters was varied. Six transformation parameters, those parameters that are used to transform the NDVI data into the time-varying boundary conditions, were also varied. Environmental driver data such as wind speed and incoming radiation were not varied in this analysis.

The full suite of input parameters was varied, rather than a subset, to ensure that unexpected parameter interactions were accounted for. Uniform distributions of parameters were chosen because there was no reliable data on means and variances with which to calculate other statistical distributions. While there have been criticisms that choosing uniform distributions might make the results of this style of analysis more sensitive to the extreme parts of the parameter range, there is also evidence that if the parameter ranges are sufficiently large yet also reasonable, any negative effects of a uniform distribution are greatly diminished (O'Neill *et al.* 1982).

Table 3.1. List of parameters varied in this experiment. The minimum and maximum values of the uniform distribution used to assign random parameters values are given. The WLEF site value typically used for point simulations is also given.

Parameter	*Definition*	*Minimum*	*Maximum*	*WLEF*
Time invariant parameters				
Z2	Canopy top height (m)	10.0	30.0	20.0
Z1	Canopy base height (m)	0.2[a]	0.8[a]	10.0
ZC	Canopy inflection height (m)	0.3[a]	0.8[a]	15.0
VCOVER	Vegetation cover fraction	0.5	1.0	0.9875
CHIL	Leaf angle distribution factor	-0.5	0.5	0.125
ROOTD	Rooting depth (m)	0.2[a]	0.9[a]	1.5
PH	½ critical leaf water potential (m)	-450.0	-50.0	-200.0
TRAN11	Green leaf transmittance (PAR)	0.0	0.1	0.05
TRAN21	Green leaf transmittance (NIR)	0.05	0.3	0.15
TRAN12	Brown leaf transmittance (PAR)	0.0	0.1	0.001
TRAN22	Brown leaf transmittance (NIR)	0.0	0.1	0.001
REF11	Green leaf reflectance (PAR)	0.02	0.2	0.07
REF21	Green leaf reflectance (NIR)	0.2	0.5	0.38
REF12	Brown leaf reflectance (PAR)	0.05	0.25	0.16
REF22	Brown leaf reflectance (NIR)	0.2	0.5	0.42
VMAX0	Rubisco velocity of sun leaf (mol m^{-2} s^{-1})	2.5E-5	15E-5	7.5E-5
EFFCON	Quantum efficiency (mol mol^{-1})	0.03	0.13	0.08
GRADM	Conductance-photosynthesis slope parameter	3.0	18.0	9.0
BINTER	Minimum stomatal conductance (mol m^{-2} s^{-1})	0.0	0.02	0.01
ATHETA	Light and rubisco coupling parameter	0.5	1.0	0.98
BTHETA	Light, rubisco and CHO sink parameter	0.5	1.0	0.95
TRDA	Respiration temperature response (K^{-1})	0.1	1.5	1.3
TRDM	Respiration inhibition ½-point temperature (K)	1.04[a]	1.1[a]	328.16
TROP	Respiration optimum temperature (K)	283.0	308.0	298.16
RESPCP	Leaf respiration fraction of Vmax	0.01	0.1	0.015
SLTI	Photosynthesis low temperature response (K^{-1})	0.1	1.5	0.2
SHTI	Photosynthesis high temperature response (K^{-1})	0.1	1.5	0.3
HLTI	Photosynthesis low temperature inhibition ½-point (K)	270.0	290.0	280.66
HHTI	Photosynthesis high temperature inhibition ½-point (K)	1.04[a]	1.1[a]	307.16
BEE	Soil wetness exponent	4.0	8.5	5.29
PHSAT	Soil water potential at saturation (m)	-0.05	-0.35	-0.25
SATCO	Saturated hydraulic conductivity	2.5E-6	100E-6	3.4E-6
POROS	Soil porosity	0.4	0.5	0.44
SLOPE	Cosine of mean terrain slope	0.1	0.25	0.1736
WOPT	Optimal percent of soil saturation for respiration	30.0	80.0	62.12
ZM	Skewness exponent of respiration vs. soil water	-0.2	0.5	0.362
WSAT	Respiration rate at soil water saturation coefficient	0.5	0.8	0.537
SODEP	Soil depth (m)	0.5	4.0	2.0
SOREF1	Soil reflectance (PAR)	0.01	0.4	0.11
SOREF2	Soil Reflectance (NIR)	1.1[a]	1.5[a]	0.22
Transformation Parameters				
LWidth	Leaf width (m)	0.003	0.1	0.04
LLength	Leaf length (m)	0.03	0.4	0.1
LTMAX	Maximum leaf area index	4.0	9.0	7.5
STEM	Stem area index	0.0	0.25	0.08
ND98	98th percentile NDVI (over 12 months, by biome)	0.5	1.0	0.686
ND02	2nd percentile NDVI (over 12 months, by biome)	0.0	0.1	0.034

[a] *Min/max values are multipliers for parameters that depend on others (see text).*

Time invariant parameters, which comprise the bulk of the input parameters, were for the most part independently randomized. However, in a few situations it was not physically realistic to vary parameters independently. For example, canopy base height (*Z1*) cannot be higher than canopy top height (*Z2*). In six cases parameters were made partially dependent on other parameters. For five of the parameters (*Z1*, *ROOTD, TRDM, HHTI, SOREF2*) a randomization interval was selected (see Table 3.1) and the random value (*R*) within this range was multiplied by the independent parameter (*I*) to give the new parameter value (*P*):

$$P = I * R \tag{3.1}$$

In this way, parameter *Z1* (canopy base height) is partially dependent on *Z2* (canopy top height), *ROOTD* (rooting depth) is partially dependent on *SODEP* (soil depth), *TRDM* (respiration inhibition ½ point temperature) is partially dependent on *TROP* (respiration optimum temperature), *HHTI* (photosynthesis high temperature inhibition ½ point) is partially dependent on *HLTI* (photosynthesis low temperature inhibition ½ point) and *SOREF2* (soil near infrared reflectance) is partially dependent on *SOREF1* (soil visible reflectance). In the sixth case, canopy inflection height (*ZC*) was made partially dependent on both canopy top height (*Z2*) and canopy base height (*Z1*) as follows:

$$ZC = (Z2 - Z1) * R + Z1 \tag{3.2}$$

Six parameters that are used to calculate the time varying parameters (transformation parameters) were also varied, they are: *LWIDTH* (leaf width), *LLENGTH* (leaf length), *LTMAX* (maximum leaf area index), *STEM* (stem area index), *ND98* (98^{th} percentile NDVI by biome) and *ND02* (2^{nd} percentile NDVI by biome). These six parameters are used to calculate the following time varying parameters: fractional absorbed photosynthetically active radiation (f_{PAR}), leaf area index (*ZLT*), greenness fraction (*GREEN*), roughness length (*Z0D*), displacement height (*ZP_DISP*), a coefficient for boundary layer resistance (*CC1*) and a coefficient for canopy layer resistance (*CC2*). The transformation parameters were varied, instead of independently adjusting the time varying parameters, to preserve the gross seasonal changes in the NDVI data while still altering the range of calculated values. In other words, the land surface will still have more vegetation in the summer months than in the winter, but the timing of leaf-on and leaf-off as well as the magnitude of the vegetation changes will be altered. The equations for the time varying parameters f_{PAR}, *ZLT*, *GREEN* and *GMDMU* are given in Appendix 2. These equations have previously been published in Sellers et al. (Sellers *et al.* 1996a) and Los et al. (Los *et al.* 2000) but are reprinted here to aid in interpretation. These equations use the transformation parameters *LTMAX*, *STEM*, *ND98*, and *ND02*. The transformation parameters *LWIDTH* and *LLENGTH*, along with *Z2, Z1, ZC, CHILL* and *VCOVER*, are used to calculate the aerodynamic parameters *Z0D*, *ZP_DISP, CC1* and *CC2*. The logic and equations for these parameters can be found in Sellers et al. (Sellers *et al.* 1996a).

For each set of parameters, the model was run with the same forcing data for six annual cycles to equilibrium to define soil moisture and soil temperature profiles. Reasonable soil moisture and soil temperature profiles have been shown to be important for land surface-atmosphere simulations at a variety of scales. However, they are notably difficult to initialize. This is due to the variable nature of soil texture and organic content across the landscape, the interaction with surface cover (vegetation, litter, snow) and the lack of appropriate data. In this study, modeled flux results are related to observations over a large area (at least 24 km^2) and, as in other modeling exercises, these variables are "spun up" with observed weather. The assumption is that the resultant temperature and moisture profiles should approach some equilibrium between the weather forcing and the model dynamics. Without a spin up, modeling results are sensitive to the transient. That is, if initial conditions are poorly specified for a particular model configuration, the effect will be anomalous results for an undetermined period of time as the model adjusts to the initial conditions. In this study the focus is on the parameterization of the model as opposed to dependence on initial conditions and so spin up of the model was considered beneficial. Further, by not randomly altering the initial conditions and driver data, common properties of the simulation are preserved and sensitivities to the input parameters alone are more straightforward to interpret.

The chosen six-year length of the spin up is conservative and was based on three sets of simulations. In the first case, the model was run with an isothermal (275.5K) and fully saturated soil replaced the initial soil conditions reported in chapter 2 (Table 2.4; here Table 3.2). In the second case an isothermal (275.5K) and completely dry soil replaced the initial soil conditions reported in chapter 2 (Table 2.4; here Table 3.2). The isothermal soil temperature (275.5) is the average of the two deepest soil layers in Table 2.4. In the third case, the model was initialized with soil temperature and moisture profiles that were variable but which were not spun-up (Table 3.2) and then run with three different random parameter sets. In each case, the model was run with the same forcing data for 6 annual cycles to generate initial conditions. The model was then run with the updated initial conditions and soil temperatures and soil moisture were compared between and among the simulations.

Table 3.2 Spin-up conditions for the WLEF site, 1997.

Initial Condition	*Wet Spinup*	*Dry Spinup*	*Parameter Spinup*
Canopy Temperature	272.6 K	272.6 K	272.6 K
Ground Temperature	273.4 K	273.4 K	273.4 K
Deep Soil Temperature – Level 1 (lower)	275.5 K	275.5 K	282.6 K
Deep Soil Temperature – Level 2	275.5 K	275.5 K	276.5 K
Deep Soil Temperature – Level 3	275.5 K	275.5 K	274.5 K
Deep Soil Temperature – Level 4	275.5 K	275.5 K	272.7 K
Deep Soil Temperature – Level 5	275.5 K	275.5 K	272.9 K
Deep Soil Temperature – Level 6 (upper)	275.5 K	275.5 K	275.5 K
Water on vegetation surface	0.0 m	0.0 m	0.0 m
Water on ground surface	0.0 m	0.0 m	0.0 m
Snow on vegetation surface	0.0 m	0.0 m	0.0 m
Snow on ground surface	0.004 m	0.004 m	0.004 m
Soil water percent of saturation level 1	1.0	0.0	0.33
Soil water percent of saturation level 2	1.0	0.0	0.20
Soil water percent of saturation level 3	1.0	0.0	0.21
Stomatal resistance	18083 mol m^{-2} s^{-1}	18083 mol m^{-2} s^{-1}	18083 mol m^{-2} s^{-1}
Turbulent kinetic energy	4.17	4.17	4.17

Figures 3.1 and 3.2 show the results of the spin-up simulations for soil percent saturation and soil temperature for the deepest represented soil layer, which is also the slowest to respond to surface inputs. In the first two simulations (Figure 3.1), with isothermal soils but different initial moisture contents, dry soils took longer to respond than wet soils. It took approximately seven years for the dry soil conditions to produce two consecutive years of identical soil temperature profiles. The random parameter spin-up took fewer years to spin-up to equilibrium (Figure 3.2), most likely because the original initial conditions were somewhat moist. However, since the soil characteristics are different in each parameter set, the spin-up times and eventual equilibrated soil temperature and moisture profiles are different for each parameter set. The conservative figure of six years was thus chosen so that each simulation would clearly be equilibrated even if extreme soil characteristics were encountered in the random parameter sets,

reflecting the true sensitivity of the model to the parameters rather than a transient sensitivity to the initialization of the parameter set.

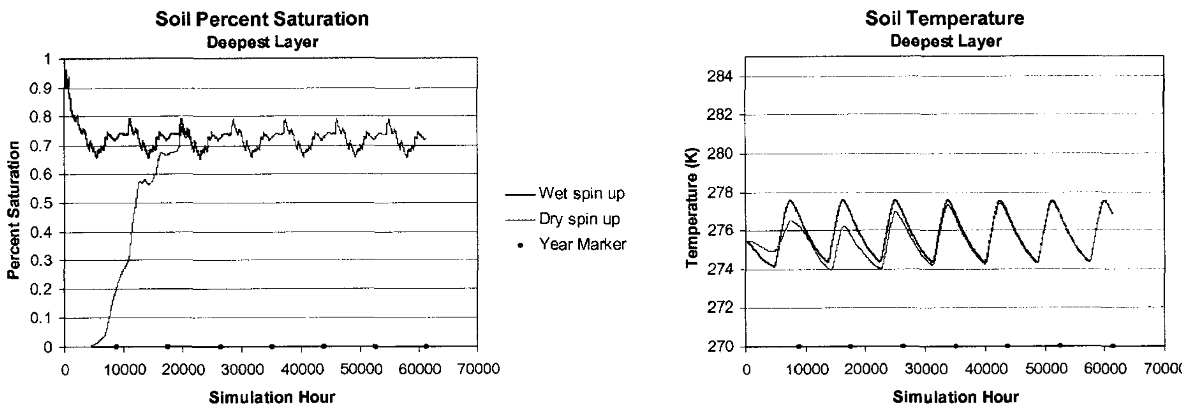


Figure 3.1. Six-year spin-up of the model plus the final model run starting from isothermal soils and either totally saturated (black) or totally dry soil (grey). The temperature of the soil for the deepest layer in dry conditions takes the longest to spin up. The red dots indicate the completion of a year.

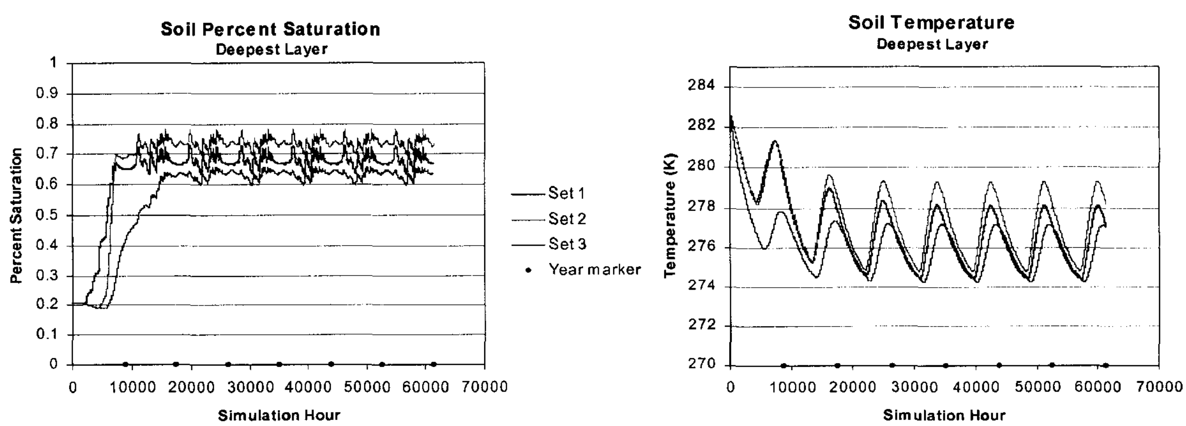


Figure 3.2. Six-year spin-up of the model plus the final model run for three different randomly generated parameter sets. Each parameter set takes a different amount of time to equilibrate and reaches a different equilibrium state.

3.3.2 *Goodness of fit criteria*

Twenty thousand sets of parameters were randomly generated using the above guidelines. The SiB2.5 model was then spun-up for six years (see above) and run for a final year (1997), 20,000 times with the randomly generated parameter sets and the meteorological driver data described in Chapter 2. In Monte Carlo simulations such as the one presented here, the preferred number of simulations to conduct to cover all possible cases is n factorial ($n!$). Clearly, for large numbers of parameters, this is prohibitive. Even ten percent of $n!$ for 46 parameters, the number of parameters analyzed here, is prohibitively large (~5.5E56). Instead, a balance between computational cost and representation of variability should be the goal. Here we have chosen to simulate 20,000 cases, which gives a satisfactory range of variation in parameter values and combinations of parameters. This is evident in scatter plots of the simulations which show good representation across the range of parameter values (e.g. Figure 3.3 below). The cost of the simulations in terms of computing time was greatly diminished by utilizing a linux cluster. The Monte Carlo runs were distributed across 20 processors on a cluster of linux computers and took in total 660 hours to run (~33 hours on each processor).

The choice of goodness of fit or likelihood criteria to measure the performance of individual simulations is important. It should clearly represent which simulations produce better results based on given criteria. Goodness of fit measures can be based on a single objective criterion such as using the root mean square error between a model prediction and an observation or on multi-objective criteria, which might include minimizing multiple single objective criteria, weighting single objective criteria and combining them or normalizing single objective criteria and combining them (Franks 1998; Gupta *et al.* 1998; Bastidas *et al.* 1999; Franks *et al.* 1999; Gupta *et al.* 1999; Meixner *et al.* 1999; Schulz *et al.* 1999; Schulz *et al.* 2001). Both single and multi-objective criteria are useful for analyzing model sensitivity, however, when well designed, multi-objective criteria add additional constraints on model performance and parameter interactions. On the other hand, in multi-objective approaches there is no unique, best solution.

Both single objective and multi-objective goodness of fit statistics were calculated for each of the 20,000 runs, comparing predicted fluxes with the observed fluxes described in Chapter 2. For the single objective criteria the root mean square difference (RMSD) between the observed (*obs*) and predicted (*est*) values of latent heat (LE), sensible heat (H) and net ecosystem exchange (NEE) was calculated as follows:

$$RMSD = \sqrt{\frac{\sum (obs - est)^2}{(n-1)}} \tag{3.3}$$

The RMSD was calculated for each month and for all 12 months combined, giving a goodness of fit statistic in units of flux.

For the multi-objective criteria a likelihood function developed by Franks (Franks 1998) was used (eq. 3.4). This likelihood function utilizes the variance of errors normalized by the minimum variance of errors and is a variation of the Nash-Sutcliffe coefficient (Nash and Sutcliffe 1970). Normalizing by the minimum variance of the errors accounts for any bias that might be caused by errors in the observations or forcing data that are particular to an individual predicted variable. This formulation also permits the multi-objective criteria to include variables with disparate units, in this case Watts and micromoles.

$$L\left(\theta \middle| \underline{y}, \sigma_\alpha^2, \sigma_\beta^2, \ldots\right) = \left(\left(\frac{\sigma_\alpha^2}{\hat{\sigma}_\alpha^2}\right)\left(\frac{\sigma_\beta^2}{\hat{\sigma}_\beta^2}\right)\ldots\right)^{-N} \tag{3.4}$$

Where:

$\theta =$ parameter set

$\underline{y} =$ observations

σ_α^2 = variance of errors for a particular variable
$\hat{\sigma}_\alpha^2$ = minimum of the variance of errors for the time period
N = scaling, shaping factor to impose stringency/leniency

SiB2.5 relies heavily on remotely sensed data for input parameters making it uniquely dependent on time varying land surface conditions. Therefore, unlike most analyses of this type, model sensitivities were evaluated monthly as well as annually. A combined likelihood statistic was calculated for LE and H (eq. 3.5) and for LE, H and NEE (eq. 3.6) for each month and for all 12 months combined, giving a unitless goodness of fit statistic.

$$L\left(\theta \middle| \underline{y}, \sigma_{LE}^2, \sigma_H^2\right) = \left(\frac{\sigma_{LE}^2}{\hat{\sigma}_{LE}^2}\right)\left(\frac{\sigma_H^2}{\hat{\sigma}_H^2}\right) \tag{3.5}$$

$$L\left(\theta \middle| \underline{y}, \sigma_{LE}^2, \sigma_H^2, \sigma_{NEE}^2\right) = \left(\frac{\sigma_{LE}^2}{\hat{\sigma}_{LE}^2}\right)\left(\frac{\sigma_H^2}{\hat{\sigma}_H^2}\right)\left(\frac{\sigma_{NEE}^2}{\hat{\sigma}_{NEE}^2}\right) \tag{3.6}$$

The scaling factor N in equation 3.4 is used to weight higher performing parameter sets more heavily when the likelihood value is used to distinguish good simulations from bad (Franks 1998). As N becomes more negative the parameter sets with the smallest variance are given more weight, resulting in a smaller set of acceptable parameter sets. In this analysis the scaling factor N was set to negative one because a percentage of the total number of parameter sets were retained as acceptable, negating the effects of N.

The two multi-objective criteria were calculated (eq. 3.5 and 3.6) to investigate the differences in model sensitivity to predictions of moisture and energy fluxes versus moisture, energy and carbon fluxes. It was hypothesized that the extra constraint that carbon fluxes impose should reduce tendencies towards equifinality. The minimum possible likelihood value for the multi-objective criteria is one and likelihood values close to one represent model runs with the least amount of error in predicted versus observed fluxes, normalized by the minimum variance. The absolute range in likelihood values is unique to each parameter/month combination.

Plots of the likelihood statistic versus parameter value for the 20,000 runs show patterns in sensitivity or equifinality for each parameter/month combination (Franks *et al.* 1997; Beven and Freer 2001; Schulz *et al.* 2001). Figure 3.3 shows an example of the parameter *ND98*, which is the biome type specific 98^{th} percentile value of NDVI used to calculate fractional absorbed photosynthetically active radiation (see Appendix 2). In Figure 3.3 the multi-objective criteria for LE and H for the months of April and July are plotted against *ND98*. Notice in April the density of points at the lower likelihood values (better simulations), particularly toward the lower range of *ND98*. This pattern indicates that in April, the value of the parameter *ND98* is important for successful model runs and that in particular values of *ND98* should be in the lower portion of it's range. In contrast, in July the scatter of points is more uniform and much less dense, indicating less sensitivity and a tendency towards equifinality (i.e. equally good simulations can be found across the full range of ND98 values). The y-axis on both plots is the same for comparison purposes, however for April this captures approximately 94% of the calculated likelihood values whereas in July this range captures only 43% of the likelihood values, the rest of the plot for July is similarly uniform.

3.3.3 *Cumulative frequency for parameter sensitivity*

To elucidate the patterns that are visible in the scatter plots, the 20,000 runs were sorted based on the likelihood estimate for each month and for all 12 months and divided into performance class groupings of 2000 (10% of the runs), similar to Franks et al. (Franks *et al.* 1997). For example, the top performance class consists of the top two thousand runs for any given parameter/month combination and the lowest two-thousand runs (ranked runs 18000-20000) represent the worst performance class. The decision to classify superior parameter sets in this way, as opposed to a combination of the shaping factor N and a likelihood value cut-off, is both objective and practical. Retaining an equal number of superior parameter sets across all combinations of month and likelihood statistics makes further analysis more straightforward and intercomparable. Choosing to work with the best ten percent of simulations also removes what may be unrealistic performance expectations.

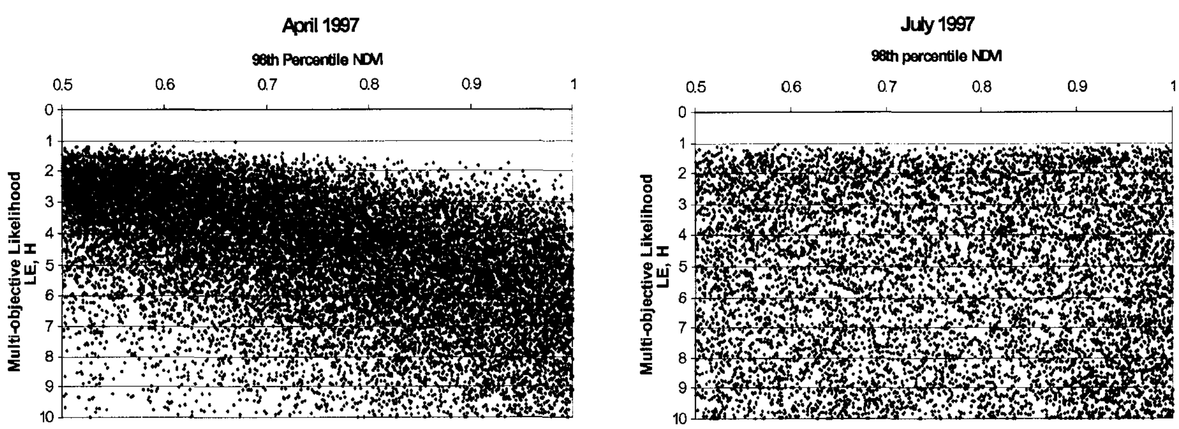
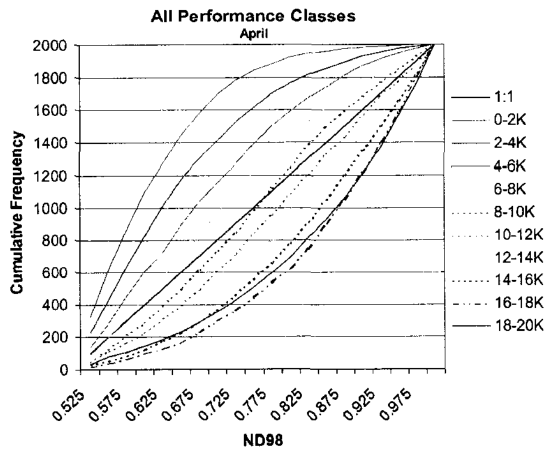


Figure 3.3 Plots of LE,H likelihood statistic against *ND98* parameter value for April and July 1997.

The cumulative frequency of parameter values in each month/performance class combination for both likelihood statistics was calculated. Cumulative frequency gives an indication of parameter sensitivity when it deviates from an expected distribution (Hornberger and Spear 1981; Franks *et al.* 1997; Bastidas *et al.* 1999; Meixner *et al.* 1999; Schulz *et al.* 1999). The range of each parameter was divided into twenty bins of uniform size. The frequency of occurrence of parameter values in each bin was calculated: for each month and for all 12 months combined, for each performance class and for both multi-objective criteria. Cumulative frequency was computed by totaling the number of occurrences in each bin range for each combination and cumulatively adding them together until 2000 was reached. Figure 3.4 shows an example of the results of this process for *ND98* for the multi-objective criteria LE and H for April and July. For both months the cumulative frequency of the top performance class is plotted in red and the worst performance class is plotted in blue. Separation between the performance classes is visible in April, indicating sensitivity. However, in July the performance classes collapse toward the 1:1 line (black; equivalent to a uniform distribution), indicating a lack of sensitivity to the parameter across the full range of *ND98*. When there is good separation between performance classes, the range of values that produce superior runs versus inferior runs is visible. The area of steepest slope indicates where the majority of parameter values fall for any given performance class.

Figure 3.5 shows the cumulative frequencies of the top performance classes across all months for the parameter *ND98* for the LE and H multi-objective criterion. Here, the most sensitive month, April, is plotted in red and the least sensitive month, July, is plotted in blue. The cumulative frequency diagram for all twelve months combined is plotted in green. This figure (3.5) shows visibly how the sensitivity of modeled results to the parameter *ND98* varies, and the importance of looking at model sensitivity through time. In this case, analyses using only one month of data or even using a full year of data in a single analysis would not provide a good assessment of the magnitude or seasonal variability of model sensitivity.



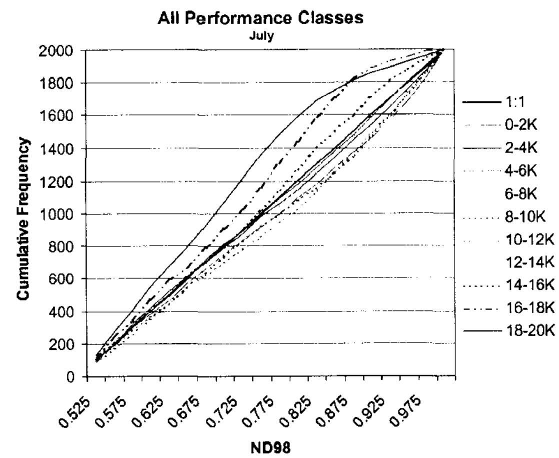


Figure 3.4 Plots of the cumulative frequency of parameter *ND98* values in each performance class for the multi-objective criterion (LE,H) for the months of April and July 1997.

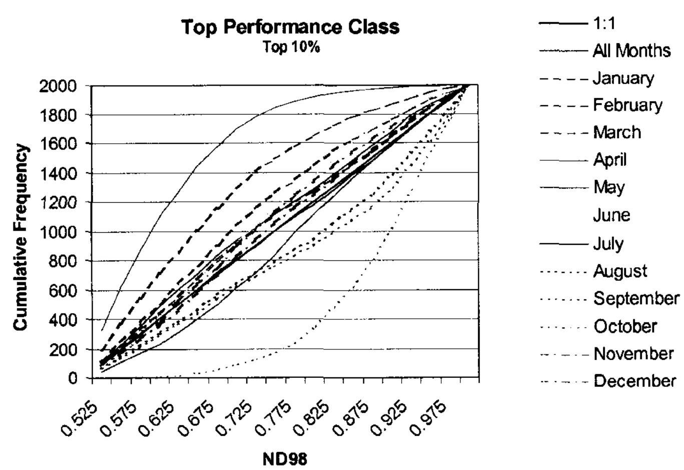


Figure 3.5 Plot of the cumulative frequency of the top performance class for the parameter *ND98*, criterion (LE,H) for each month in 1997.

3.3.4 *Parameter sensitivity test*

The top 2000 runs (top ten percent) for each multi-objective criterion for each parameter/month combination were further evaluated, 1,196 combinations in all (46 parameters x 12 months + annual x 2 criteria). A Kolmogorov-Smirnov test for goodness of fit (Sokal and Rohlf 1981) was performed for each cumulative frequency distribution for every month/parameter/criterion combination. The Kolmogorov–Smirnov test (K-S test) compares cumulative frequency distributions for populations that don't conform to standard probability models (non-parametric) by testing the difference between an observed (F) and an expected $(\hat{F})$ distribution. As we have seen above (Figure 3.5) when any value of a parameter has an equal chance of producing a good simulation, the shape of the cumulative frequency of that parameter is a straight line. The hypothesis that is being tested with the Kolmogorov-Smirnov test is that there is no difference from the expected distribution (1:1), i.e. the specification of a given parameter has no influence on the results. This is an extrinsic hypothesis because it is external to the data being tested and was formed from a theoretical assumption and therefore the critical value used here is for a one-sample Kolmogorov-Smirnov statistic. The Kolmogorov-Smirnov test statistic has been successfully used in sensitivity tests (Hornberger and Spear 1981; Bastidas *et al.* 1999; Meixner *et al.* 1999). The Kolmogorov-Smirnov test statistic is sensitive to the central tendency and distribution of a given parameter (Hornberger and Spear 1981). It is flexible because it can be used for one and two sample tests and for extrinsic (external to the data) and intrinsic (generated from the data) hypotheses. Another desirable property of the K-S test is that significance can be ranked based on the K-S test statistic value.

In practice, the Kolmogorov-Smirnov (K-S) test for goodness of fit tests the largest absolute difference between the observed cumulative frequency and the expected cumulative frequency against critical values, allowing the null hypothesis to be rejected at a particular significance level (e.g. α=0.01). The Kolmogorov-Smirnov test statistic is the largest absolute difference between the observed cumulative frequency minus 0.5 and the expected cumulative frequency (MaxAbs) divided by the sample size (n) (eq. 3.7) (Sokal and Rohlf 1995).

$$D_{\max} = \frac{MaxAbs}{n} \tag{3.7}$$

Table 3.3 shows the cumulative frequency of the theoretical distribution (1:1) and the cumulative frequencies for the top performance class of *ND98* for April and July 1997, LE,H criterion. The last two columns show the absolute difference between the theoretical distribution and the observed distributions. The maximum difference, which is highlighted, is used to calculate the K-S statistic.

Table 3.3 Calculation of the K-S test statistic.

ND98 Bins	**(a) Theoretical Distribution**	**(b) April ND98 Distribution - 0.5**	**(c) July ND98 Distribution – 0.5**	**Absolute value (a-b)**	**Absolute value (a-c)**
0.525	100	325.5	84.5	225.5	15.5
0.55	200	615.5	185.5	415.5	14.5
0.575	300	864.5	293.5	564.5	6.5
0.6	400	1093.5	396.5	693.5	3.5
0.625	500	1268.5	493.5	768.5	6.5
0.65	600	1445.5	592.5	845.5	7.5
0.675	700	1576.5	717.5	876.5	17.5
0.7	800	1695.5	809.5	**895.5**	9.5
0.725	900	1782.5	892.5	882.5	7.5
0.75	1000	1842.5	991.5	842.5	8.5
0.775	1100	1884.5	1093.5	784.5	6.5
0.8	1200	1921.5	1173.5	721.5	**26.5**
0.825	1300	1945.5	1274.5	645.5	25.5
0.85	1400	1959.5	1381.5	559.5	18.5
0.875	1500	1976.5	1475.5	476.5	24.5
0.9	1600	1981.5	1582.5	381.5	17.5
0.925	1700	1988.5	1692.5	288.5	7.5
0.95	1800	1993.5	1795.5	193.5	4.5
0.975	1900	1997.5	1890.5	97.5	9.5
1	2000	1999.5	1999.5	0.5	0.5

The largest absolute difference between columns (a) and (b) is 896.5, the largest absolute difference between columns (a) and (c) is 26.5. The D_{max} values are then:

April D_{max} = (896.5)/2000 = 0.448

July D_{max} = (26.5)/2000 = 0.013

The critical value of D at α=0.01, n=2000 is (Sokal and Rohlf 1995):

$$g_\alpha = \sqrt{\frac{-\ln(\alpha/2)}{2n}} + \frac{1}{2n} = 0.03614 \tag{3.8}$$

For April the null hypothesis, that the observed distribution is similar to the expected distribution, is rejected since 0.448 is greater than 0.03614. For July the null hypothesis is accepted since 0.013 is less than 0.03614. Therefore simulations of LE and H for the month of April are sensitive to the specification of the parameter *ND98* whereas simulations of the same quantities in July are not. The total number of sensitive parameters were calculated in this way for both multi-objective criteria (LE,H and NEE,LE,H) for all month/parameter combinations.

3.4 Results

3.4.1 *Minimum RMSD errors*

The root mean square difference (RMSD; eq. 3.3) was calculated between simulations and observations of net ecosystem exchange, latent heat flux and sensible heat flux for each month and all months combined for all 20,000 parameter sets. An interesting result was the minimum level of error between observations and simulations that could not be surpassed despite the many parameter combinations and thousands of simulations. Table 3.4 reports the minimum RMSD errors for each simulated flux. The magnitude of the minimum error varies throughout the year. For net ecosystem exchange and latent heat flux the largest errors are during the growing season: June, July, August and September. For sensible heat flux, the largest errors are during the shoulder seasons: April, May and October as well as in January. In general, the largest errors occur in times of highest flux.

Table 3.4. Minimum root mean square difference between observations and simulations across all 20,000 simulations.

	Minimum Root Mean Square Difference		
Month	**Net Ecosystem Exchange (μmol m^{-2} s^{-1})**	**Latent Heat Flux (W m^{-2})**	**Sensible Heat Flux (W m^{-2})**
All	2.68	42.04	36.60
January	0.66	8.05	41.25
February	0.85	15.00	30.92
March	0.65	19.81	31.08
April	1.27	23.20	39.36
May	1.73	37.85	42.69
June	4.56	62.42	35.30
July	4.44	70.38	28.82
August	4.59	64.21	26.76
September	3.09	52.50	26.32
October	1.93	32.58	34.16
November	1.19	14.23	24.76
December	0.46	11.00	17.82

The minimum possible error reflects a level of variability (or noise) in the observations, structural problems with the model and/or mismatch between the model and observations that cannot be overcome by varying parameters. The variability in the observations may be related to measurement errors; surveys of the FLUXNET data suggest there is a mean energy imbalance in the measurements of approximately 20% (Wilson *et al.* 2002) and other research suggests that errors in NEE are around 7-12% (Law *et al.* 2002; Baldocchi 2003). It is also possible that the variability in the measurements stems from the heterogeneous nature of the surface and the changing footprint of the measurements, though the data processing techniques were designed to minimize this. Because it is calculated here on a monthly and annual basis, the minimum possible error reflects a minimum level of uncertainty that can be associated with predictions made on those time scales. If the RMSD errors were calculated on a different time scale, such as weekly or daily, the minimum possible error would be a reflection of the average uncertainty at that time scale. This is potentially a very helpful number that might be used, for example, as a prior uncertainty in Bayesian synthesis inversion studies.

3.4.2 *Kolmogorov-Smirnov test statistics: General results*

The Kolmogorov-Smirnov (K-S) test statistics for each month/parameter combination for both the LE,H and the NEE,LE,H multi-objective criteria are reported in Tables 3.5 and 3.6 respectively. Values of the test statistic which are less than the p=0.01 significance level are shaded yellow, values which are greater than or equal to the p=0.01 significance level are shaded in orange. Parameters are considered influential if they are significant at the p=0.01 level, however the higher the K-S test statistic, the more sensitive the model is to the specification of that parameter. Appendix 3 contains a companion table (Table A3.1) that combines the information in Tables 3.5 and 3.6 to show which parameters are influential according to the LE,H criterion, the NEE,LE,H criterion, both criteria or neither. Parameters are grouped by function to elucidate patterns of sensitivity and insensitivity.

It is apparent, looking at these tables collectively, that model sensitivity varies by criterion and by the time period considered. For example, the column 'All Months' represents the sensitivity of SiB2.5 to any given parameter considering the entire 12 months of simulation and observations. There are many instances, however, where the model is sensitive to parameters in specific months but not when a whole year is considered. Thus, if only that one measure was considered, a significant amount of information would be lost. It is also clear that the types of influential parameters vary by month. In the growing season, when NEE and LE are tightly coupled, both multi-objective criterion (which represent energy balance and energy balance plus the addition of NEE) produce similar results for the photosynthesis parameters. In the dormant season, the additional constraint by the photosynthesis parameters does not impact the prediction of LE and H much.

Table 3.5 Kolmogorov-Smirnov test statistics for the LE,H multi-objective criterion. Parameters are arranged by usage in the model and shaded to reflect whether the parameter was found influential. Parameters significant to modeled results at the p=0.01 significance level are shaded orange. Those that are not significant are shaded yellow.

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All
Aerodynamic													
lwidth	0.03575	[illegible]	0.02025	[illegible]	0.14025	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02825	[illegible]
llength	0.02775	0.02925	0.02075	[illegible]	0.01625	0.01575	0.00725	0.01475	0.01975	0.02875	[illegible]	0.01475	0.01575
z2	[illegible]	[illegible]	[illegible]	0.03375	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
z1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
zc	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Remote Sensing													
nd98	[illegible]	[illegible]	[illegible]	[illegible]	0.10275	[illegible]	0.01325	[illegible]	[illegible]	[illegible]	0.03075	[illegible]	[illegible]
nd02	0.01775	[illegible]	[illegible]	[illegible]	0.01275	0.01375	0.01275	0.02375	0.02225	0.03425	0.02425	0.01975	[illegible]
ltmax	0.02425	[illegible]	[illegible]	[illegible]	0.01275	[illegible]	0.01675	[illegible]	[illegible]	[illegible]	0.01825	[illegible]	[illegible]
stem	0.02225	[illegible]	[illegible]	[illegible]	[illegible]	0.01275	0.02275	0.03575	[illegible]	0.02375	0.01625	[illegible]	[illegible]
Energy Balance													
soref1	0.02525	0.01975	0.01825	[illegible]	[illegible]	0.03275	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
soref2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
vcover	[illegible]	0.01225	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02525	[illegible]	0.02425	[illegible]
chil	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02525	0.03025	[illegible]	0.03025	[illegible]	[illegible]	0.02475	[illegible]
ref11	0.01925	0.01475	0.01475	0.02075	[illegible]	0.00975	[illegible]	0.03275	[illegible]	0.03475	0.01125	0.03025	0.01225
ref21	0.00725	0.03075	[illegible]	[illegible]	[illegible]	0.01375	[illegible]	[illegible]	[illegible]	[illegible]	0.02975	[illegible]	0.01325
ref12	0.02775	0.02175	0.02175	0.00725	0.01475	0.01175	0.01325	0.01725	0.01075	0.02675	0.02025	0.01625	0.02425
ref22	0.02525	0.02125	[illegible]	0.01175	0.00825	0.01325	0.01975	0.01425	0.02075	0.03075	0.00875	0.02425	0.02725
tran11	0.01025	0.01075	0.02275	0.02725	0.02725	0.01375	0.01525	0.00825	0.01825	0.01325	0.01275	0.03025	0.02725
tran21	0.02775	[illegible]	[illegible]	[illegible]	[illegible]	0.01725	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.01925
tran12	0.02225	0.01475	0.01775	0.01025	0.01375	0.01275	0.01075	0.01425	0.01375	0.01675	0.01525	0.01625	0.01625
tran22	0.01175	0.01875	0.02775	0.00775	0.01625	0.02475	0.00875	0.00675	0.01725	0.02775	0.01575	0.02175	0.01575
Soil Moisture Availability													
bee	[illegible]	[illegible]	0.02425	[illegible]	[illegible]	0.01575	0.03525	0.02975	[illegible]	[illegible]	[illegible]	[illegible]	0.02675
phsat	0.02725	0.01325	0.01375	0.02925	0.02025	0.01725	0.01325	0.01225	0.01525	0.01525	[illegible]	0.02125	0.01575
satco	0.03075	0.01975	0.01475	[illegible]	[illegible]	0.03325	0.01675	0.02175	0.02225	[illegible]	[illegible]	[illegible]	[illegible]
poros	[illegible]	0.01975	0.02225	0.03575	0.02325	[illegible]	0.02725	0.02875	0.01775	[illegible]	[illegible]	[illegible]	0.03075
sodep	[illegible]	[illegible]	0.02775	[illegible]	[illegible]	[illegible]	0.01825	0.01325	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
rootd	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Photosynthesis													
vmax0	0.01525	0.01325	0.01725	0.14075	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.01575	0.03425	0.01975	[illegible]
effcon	0.02725	0.01825	0.01775	0.02225	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.01375	0.01425	0.02375	[illegible]
gradm	0.01775	0.01425	0.02375	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.01875	[illegible]
binter	0.02625	0.02275	0.01175	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02025	0.02575	[illegible]
atheta	0.01125	0.01825	0.01375	0.02825	[illegible]	0.02975	[illegible]	[illegible]	[illegible]	0.01775	0.01925	0.01525	0.03325
btheta	0.02075	0.01375	0.02475	0.01675	0.02875	0.02525	[illegible]	[illegible]	[illegible]	0.02725	0.00875	0.02575	0.02625
hiti	0.01175	0.01675	0.02575	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02275	0.01075	0.02325	[illegible]
hhti	[illegible]	0.14775	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
siti	0.01275	0.02425	0.01025	0.01125	0.03075	0.02175	0.01375	0.01025	0.01725	0.02075	0.02225	0.01725	0.01375
shti	0.01425	0.02375	0.01125	0.02175	0.03175	[illegible]	0.01225	0.02225	0.02225	0.02325	0.01475	0.01775	0.00875
ph	0.00875	0.01525	0.00975	0.03175	0.01425	0.02625	[illegible]	[illegible]	[illegible]	0.01425	0.01875	0.01625	[illegible]
Autotrophic respiration													
respcp	0.02375	0.02275	0.01575	[illegible]	0.03475	0.10775	[illegible]	[illegible]	[illegible]	0.02375	0.01725	0.02675	[illegible]
trda	0.01425	0.01175	0.00875	0.02075	0.01675	0.01125	0.01825	0.01625	0.01525	0.01175	0.01625	0.01475	0.01425
trdm	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
trop	0.02125	0.01675	0.00925	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.03075	0.03125	[illegible]	[illegible]
Heterotrophic respiration													
wopt	0.01875	0.01975	0.01325	0.02175	0.01325	0.01275	0.01325	0.01275	0.01325	0.01875	0.02475	0.01925	0.01425
zm	0.01375	0.01175	0.01225	0.01675	0.02075	0.00925	0.00625	0.01225	0.01125	0.01875	0.00925	0.02725	0.01475
wsat	0.01675	0.01275	0.01575	0.02275	0.01275	0.01475	0.01475	0.01025	0.01775	0.01425	0.01675	0.01175	0.01525

not significant at p=0.01

Table 3.6 Kolmogorov-Smirnov test statistics for the NEE,LE,H multi-objective criterion. Parameters are arranged by usage in the model and shaded to reflect whether the parameter was found influential. Parameters significant to modeled results at the p=0.01 significance level are shaded orange. Those that are not significant are shaded yellow.

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All
Aerodynamic													
lwidth	0.03275	0.04225	0.01625	[illegible]	0.05725	0.02225	0.02925	0.10775	[illegible]	[illegible]	[illegible]	0.01575	[illegible]
llength	0.02825	0.01525	0.01725	0.01875	0.02925	0.01875	0.01525	0.01375	0.00925	0.01675	[illegible]	0.01475	0.01525
z2	0.34175	[illegible]	[illegible]	0.02525	[illegible]	0.01325	0.03525	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
z1	0.26775	[illegible]	[illegible]	0.16125	[illegible]	0.24725	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
zc	[illegible]	[illegible]	0.21075	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Remote Sensing													
nd98	0.07725	[illegible]	[illegible]	[illegible]	[illegible]	0.10875	[illegible]	0.08775	[illegible]	[illegible]	[illegible]	0.04475	[illegible]
nd02	0.00825	[illegible]	[illegible]	[illegible]	0.02425	0.01325	0.01125	0.01325	0.01125	0.00975	0.02325	0.01525	0.02225
ltmax	0.03075	[illegible]	[illegible]	[illegible]	[illegible]	0.02375	0.02275	[illegible]	0.03325	[illegible]	0.01875	0.03125	0.02375
stem	0.02525	[illegible]	[illegible]	[illegible]	[illegible]	0.01525	0.01975	0.02975	[illegible]	[illegible]	0.01275	[illegible]	0.02525
Energy Balance													
soref1	0.03525	0.01275	0.02325	[illegible]	[illegible]	0.02925	[illegible]	[illegible]	[illegible]	0.02575	[illegible]	[illegible]	0.02475
soref2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
vcover	[illegible]	0.01875	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.03525	[illegible]
chil	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02975	0.02425	[illegible]	0.02525	[illegible]	[illegible]	0.02975	[illegible]
ref11	0.02425	0.01425	0.02475	0.02425	0.02775	0.01775	0.03225	0.03225	[illegible]	0.00875	0.01125	0.03175	0.01475
ref21	0.01325	0.02175	0.03525	[illegible]	[illegible]	0.01575	[illegible]	[illegible]	[illegible]	0.02925	0.02325	0.03125	0.02275
ref12	0.01025	0.01025	0.02675	0.01275	0.01075	0.02075	0.02475	0.01625	0.01325	0.02775	0.01275	0.01075	0.02575
ref22	0.01225	0.02575	0.04175	0.00975	0.01125	0.00975	0.02175	0.00875	0.01025	0.02575	0.01025	0.01325	0.02475
tran11	0.02475	0.02675	0.02575	0.02925	0.01575	0.00875	0.01225	0.01475	0.02275	0.01575	0.01575	0.01525	0.01075
tran21	0.01175	[illegible]	[illegible]	[illegible]	[illegible]	0.02025	[illegible]	0.03475	[illegible]	0.02825	0.01925	0.02575	0.03025
tran12	0.02325	0.01875	0.02525	0.01325	0.01825	0.00975	0.00975	0.01375	0.01875	0.00875	0.01825	0.01225	0.01025
tran22	0.01725	0.00975	0.01775	0.00725	0.01425	[illegible]	0.01675	0.01075	0.01075	0.01525	0.01475	0.00975	0.02225
Soil Moisture Availability													
bee	[illegible]	0.04475	0.01875	[illegible]	[illegible]	0.02025	0.03075	0.02575	0.03575	[illegible]	[illegible]	[illegible]	0.01875
phsat	0.02025	0.01825	0.01175	0.02225	0.00975	0.01475	0.01375	0.02075	0.00875	0.01425	0.03325	0.01625	0.01225
satco	0.02025	0.02175	0.00975	[illegible]	[illegible]	0.02325	0.01725	0.01275	0.01175	[illegible]	[illegible]	0.02225	0.03125
poros	[illegible]	0.00725	0.01575	0.02625	0.01625	[illegible]	0.02675	0.03425	0.02125	[illegible]	0.02475	0.02975	0.03575
sodep	0.05125	0.01575	0.03575	[illegible]	0.02175	[illegible]	0.01625	0.01375	0.02275	[illegible]	[illegible]	[illegible]	[illegible]
rootd	[illegible]	0.28175	0.31075	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Photosynthesis													
vmax0	0.03925	0.02375	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
effcon	0.02925	0.06275	0.02825	0.01875	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02875	0.02925	[illegible]
gradm	0.02475	0.02425	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.03525	0.02175	[illegible]
binter	0.02275	0.02025	0.02475	[illegible]	0.01575	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.02475	0.02325	[illegible]
atheta	0.00875	0.02475	0.01475	0.02875	[illegible]	0.02825	[illegible]	[illegible]	[illegible]	0.01825	0.01325	0.01325	[illegible]
btheta	0.03475	[illegible]	0.01325	[illegible]	[illegible]	0.03025	[illegible]	[illegible]	[illegible]	[illegible]	0.00975	0.01175	0.03025
hlti	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
hhti	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
slti	0.01775	0.01575	0.00475	0.02925	[illegible]	0.01275	0.00825	0.01175	0.01475	[illegible]	[illegible]	0.03075	0.02675
shti	0.01275	0.00775	0.00775	[illegible]	[illegible]	[illegible]	0.01425	0.01525	0.01575	[illegible]	0.02375	0.01725	0.01175
ph	0.01125	0.01025	0.02625	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.00725	0.02925	[illegible]
Autotrophic respiration													
respcp	0.14225	0.17375	0.11425	0.01775	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
trda	0.00975	0.01225	0.02425	0.02275	0.01075	0.01875	0.01825	0.01775	0.01025	0.01675	0.01475	0.01725	0.02075
trdm	0.16725	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
trop	0.10225	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Heterotrophic respiration													
wopt	0.02025	0.01375	0.01475	0.02525	0.01325	0.01475	0.00875	0.01425	0.00925	0.01775	0.02425	0.00925	0.01775
zm	0.01975	0.01675	0.00775	0.01575	0.01175	0.01775	0.01425	0.01475	0.02575	0.01075	0.01175	0.01325	0.00975
wsat	0.03475	0.02275	0.01725	[illegible]	0.01775	0.01225	0.01475	0.01475	0.01525	0.02525	0.03025	0.02425	0.01875

not significant   at p=0.01

The total number of influential parameters varies through time for the two multi-objective criteria, from a low of 30 percent of the parameters (LE,H criterion in January) to a high of 63 percent of the parameters (NEE,LE,H criterion in May) (Figure 3.6). At any given time, model results are independent of nearly half the input parameters. This is not to say that these parameters are not important at all, but that their specification has much less of an influence on modeled results than other parameters. In general, the model is sensitive to more parameters in the spring and fall than in the summer and winter. Which criterion produces the highest number of influential parameters also changes through time: in the summer months the LE,H criterion has the highest number of influential parameters whereas during the rest of the year the NEE,LE,H criterion does.

The opposite pattern is seen in the number of parameter sets common to both criteria for the top 2000 runs (Figure 3.7). The highest number of common runs occurs in the growing season and in the winter, the lowest number of common runs occurs in the spring and fall. A greater proportion of common runs indicates that the processes controlling NEE, LE and H are more coupled during those time periods. The pronounced seasonal difference indicates that the degree of coupling changes through time.

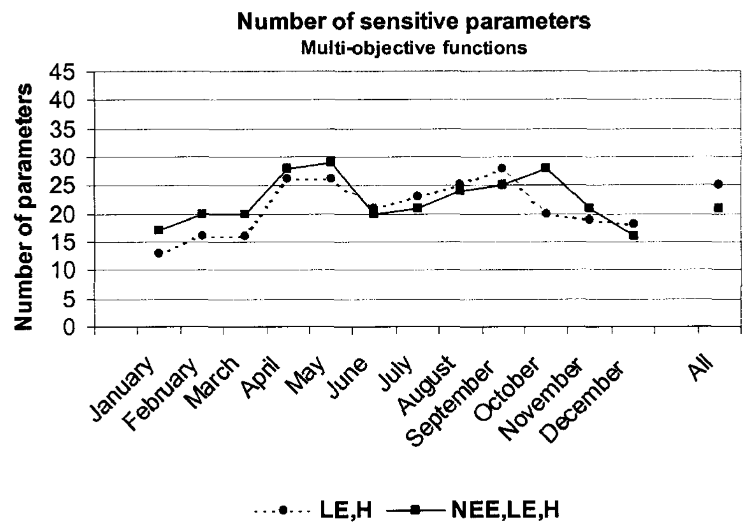


Figure 3.6. Variation in the number of sensitive parameters through time for both multi-objective criteria.

This result demonstrates the difference between NEE in the growing season, which is dominated by photosynthesis, and NEE in the dormant months, which is dominated by heterotrophic respiration. CO_2 enters and water vapor exits plants through stomatal apertures, thus the regulation of the assimilation of carbon and water loss by the plant are inextricably linked. In SiB2.5 this relationship is expressed via the Ball-Berry equation (Ball *et al.* 1987; Sellers *et al.* 1992a). In SiB2.5, heterotrophic respiration is dependant on temperature, moisture, soil texture and the annual uptake of CO_2 by the canopy (Raich and Nadelhoffer 1989; Denning *et al.* 1996a) and as such is less singularly coupled to sensible heat flux. Since photosynthesis is more tightly coupled to latent heat flux then heterotrophic respiration is to sensible heat flux, the number of common parameter sets is greatest in the growing season when latent heat is the more dominant flux as opposed to the dormant season when sensible heat is more dominant. Further, fewer common parameter sets indicates that the addition of NEE to the multi-objective criterion adds a further constraint on overall model performance, which we see expressed in the number of influential parameters (Figure 3.6). The effect of adding NEE as a performance constraint is particularly strong in May and October when the soil was thawed but there were no leaves on the deciduous vegetation.

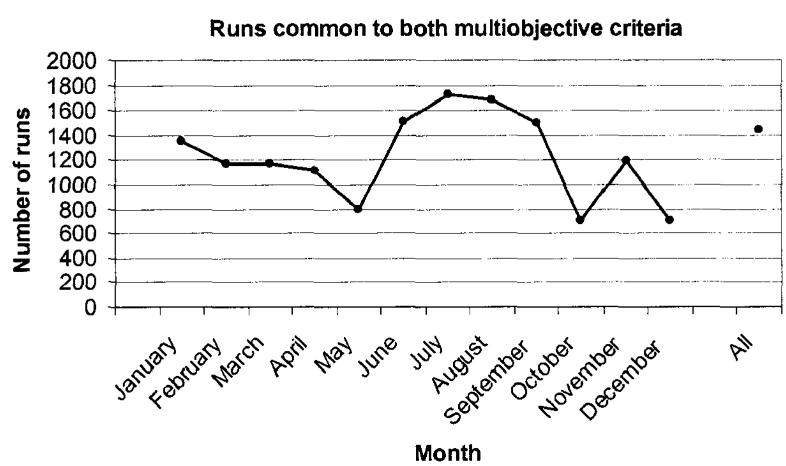


Figure 3.7. Number of parameter sets common to both multi-objective criteria for the top 2000 runs. A larger proportion of common runs indicates a stronger degree of coupling between energy balance and net ecosystem exchange.

As mentioned above, there are a number of parameters to which the model exhibits little or no sensitivity. There are six parameters that are never influential: brown leaf reflectance in the visible (*REF1,2*), green leaf transmittance in the visible (*TRAN1,1*), brown leaf transmittance in the visible (*TRAN1,2*), autotrophic respiration temperature response (*TRDA*), optimal percent of soil saturation for respiration (*WOPT*) and the skewness exponent of respiration versus soil water (*ZM*). There are also a number of parameters (7) that are only marginally influential, appearing as significant only in one or two months, and then just barely. These marginally influential parameters are: leaf length (*LLENGTH*), green leaf reflectance in the visible (*REF1,1*), brown leaf reflectance in the near infrared (*REF2,2*), brown leaf transmittance in the near infrared (*TRAN2,2*), soil water potential at saturation (*PHSAT*), soil porosity (*POROS*) and a coefficient for respiration rate at soil water saturation (*WSAT*). It is possible that some of these soil parameters might have been more influential during a dry year at WLEF.

Of the remaining parameters, some exhibit a remarkably strong influence for brief periods of time. It appears (and this will be discussed further below) that these strong represent specific cases of extreme sensitivity related to underlying problems in the modeling environment. Continuous levels of sensitivity appear to represent the more general case of model sensitivity to reasonable parameterization. Three main trends are: 1) more sensitivity during the growing season, 2) more sensitivity in winter and 3) more sensitivity during the shoulder seasons (which are the transitional time periods just before, during and after leaf-out and leaf-fall).

The following sections, arranged by thematic group, examine in more detail the annual patterns of sensitivity of some of the most influential parameters. Appendix 4 contains graphs of the Kolmogorov-Smirnov test statistics plotted through time and arranged by the same thematic groups. These plots represent the information in Tables

3.5 and 3.6 plotted in graphic form, are visually very instructive and serve as a companion to the following sections.

3.4.3 *Aerodynamic parameters*

The "aerodynamic" model parameters are used to calculate aerodynamic resistances, both within SiB2.5 as well as offline in the generation of other parameters. The aerodynamic resistances regulate the fluxes of heat, moisture and CO_2 between the leaf and soil to the canopy air space and the atmosphere. Canopy base height (*Z1)* and canopy top height (*Z2)* are also used to calculate the mass of snow on the canopy, which is then used in the two-stream radiation approximation that determines the amount of energy available for heat fluxes, to calculate the amount of water on the canopy available for evaporation and to update resistances for the presence of snow. The model is sensitive to the specification of *Z1* and canopy inflection height (*ZC*) for all months and for *Z2* in nearly all months for both criteria. Since the aerodynamic resistances regulate all of the fluxes, it might be expected that the model would be sensitive to the parameters that determine their specification year-round. It is particularly sensitive, however, in the winter months, October through February. An examination of the frequency histograms for the parameters *Z1* and *Z2* using the LE,H criterion shows that in the winter months, parameter sets which produce the best runs have both shorter canopies (Figures 3.8a,b) and smaller crown depths (Figure 3.9a,b). Clearly, it is not reasonable to expect canopy height to change a great deal during the year in a mixed deciduous and coniferous forest. In fact, the histogram of canopy top height for January, which is very asymmetric and heavily weighted towards the lower range of possibilities, indicates that an even larger range of values could have been explored. However such a short canopy would not have been appropriate given what we know of actual forest height in the WLEF region, which is around 20m (http://cheas.psu.edu). This result indicates a potential problem in winter with respect to either aerodynamic resistances or radiative transfer in the presence of snow.

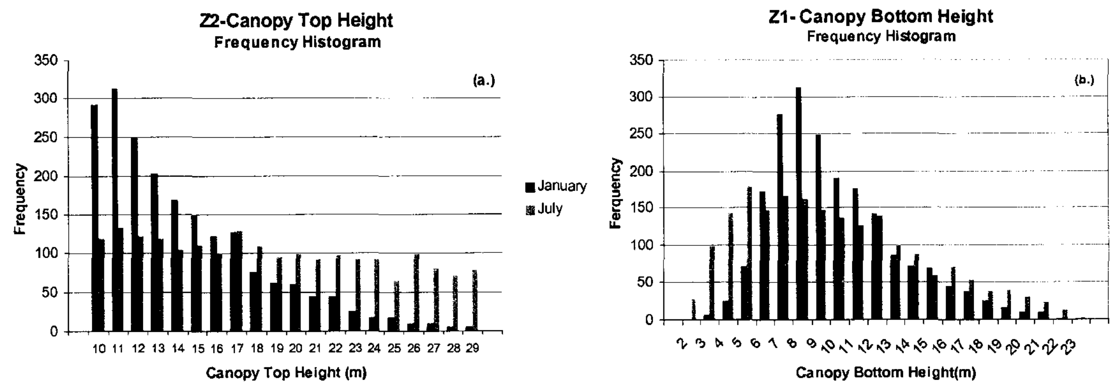


Figure 3.8. Frequency histograms of fitted parameter values in the "best" 2000 simulation using the LE,H criterion. (a.) *Z2* and (b.) *Z1* for the months January (black) and July (grey). January is representative of the winter months, July is representative of the summer months.

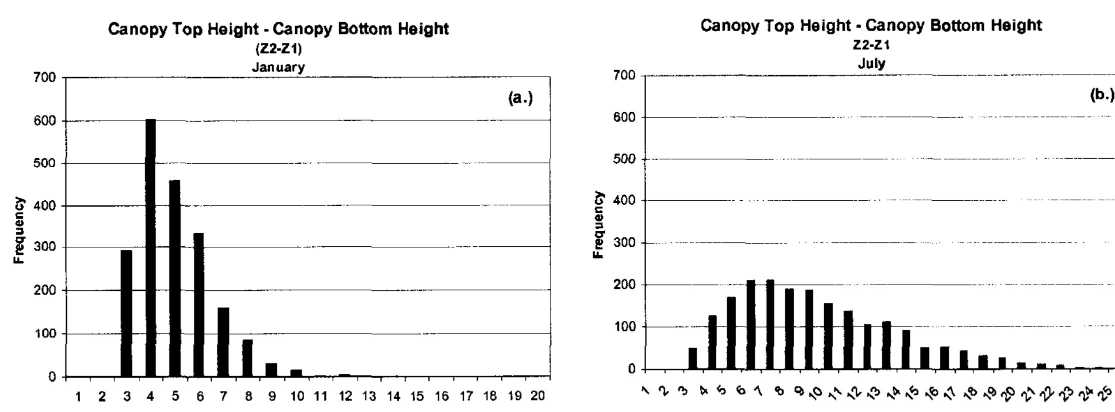


Figure 3.9. Frequency histograms of the difference between fitted parameter values *Z2* and *Z1* for (a.) January and (b.) July, shown in Figure 3.8.

3.4.4 *Remote sensing parameters*

The 'remote sensing' parameters are used offline to generate other time varying parameters. Indirectly, the parameters *ND98* (98^{th} percentile NDVI), *ND02* (2^{nd} percentile

NDVI), *LTMAX* (maximum leaf area) and *STEM* (stem area index) are all used in the calculation of the available energy and in the calculation of net assimilation of carbon via such quantities as f_{PAR} and leaf area index (Appendix 2). Shifts in these four parameters affect both the amount of f_{PAR} as well as the amount of leaf area.

ND98 is calculated by taking an annual cycle of NDVI for an entire biome and calculating the 98th percentile value of NDVI. Calculations of f_{PAR} from NDVI are then scaled between 0 and 0.95, where ND98 is equal to an f_{PAR} of 0.95 . SiB2.5 shows the greatest sensitivity to *ND98* in April and May and then again in October, the shoulder seasons. Cumulative frequency histograms of the parameter *ND98* for both criteria reflect adjustment of the parameterization to changing sensitivities through time. Lower values of *ND98* effectively increase leaf area and f_{PAR}. Higher values of *ND98* effectively decrease leaf area and f_{PAR}. For the LE,H criterion and the parameter *ND98*, the best performing values in April are in the lower part of its range while the best performing values in October are in the upper part of its range (Figure 3.10a). For the NEE,LE,H criterion, the situation is similar to the LE,H criterion, though less extreme, in April and October (Figure 3.10b). In May however, unlike in the LE,H criterion, the best performing values for the parameter *ND98* are also in the upper part of its range.

In chapter two we saw that in April soils are beginning to thaw but leaf out does not occur until the end of May. In the standard model configuration for WLEF, monthly diurnal average plots show that SiB2.5 tends to overestimate net ecosystem exchange during the daytime in May (Appendix 1). A possible explanation for the overestimation was a premature greening up of the canopy due to temporal interpolation between monthly mean NDVI composites. Thus the shift to lower f_{PAR} and leaf area index in May is understandable in that it would lead to a reduction in net ecosystem exchange and would reflect the late May leaf out better. A similar explanation is possible for October. Since leaf fall actually occurred early in the month, a reduction in leaf area would better match the condition of the actual canopy. Given that the sensitivity is much greater in the LE,H criterion, the implication is that the adjustment would benefit energy balance. A higher temporal resolution for the time varying parameters in SiB2.5 would likely solve the problem in May and October. Instead of linearly ramping quantities such as LAI from the center points of each month, shorter compositing periods of NDVI could be used that would more directly relate to the rapidly changing surface properties at such sensitive times of the year. However, if clouds are a problem in the study area, a shorter compositing period might not be possible. April is more difficult to interpret. Increasing f_{PAR} and leaf area in April would decrease the albedo of the surface, increasing available energy. The sensitivity is expressed most strongly in the LE,H criterion, again implying a stronger relationship with energy balance.

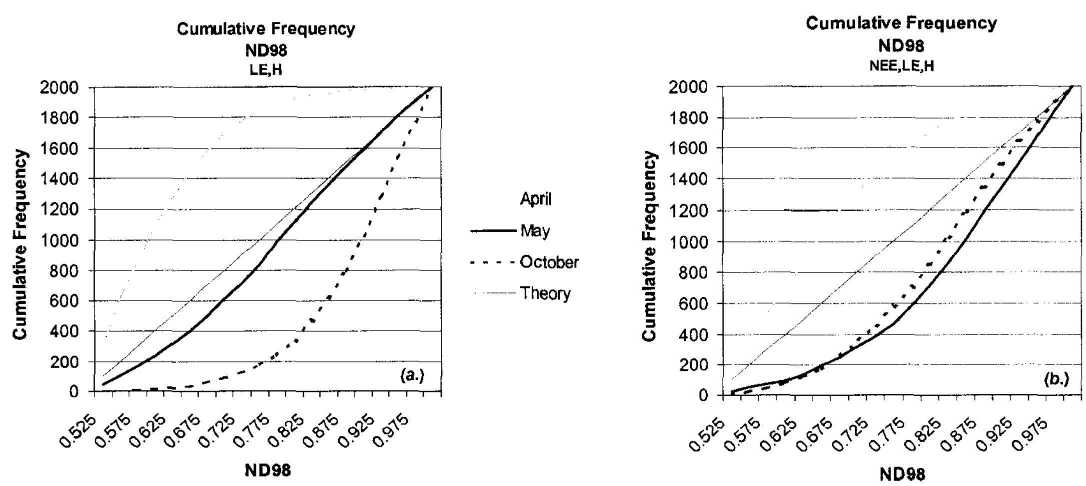


Figure 3.10. Cumulative frequency diagrams of the top 2000 fitted values for the parameter *ND98* for the months April, May and October. **(a.)** LE,H criterion and **(b.)** NEE,LE,H criterion.

3.4.5 *Energy balance parameters*

Of the energy balance parameters, SiB2.5 is most sensitive to soil reflectance (*SOREF1* and *SOREF2*). In standard SiB2.5 simulations, soil reflectance is assigned by biome, meaning that all soils for a particular type of biome are considered to have the same reflectance properties. These reflectance properties do not change with soil texture, time or the status of the soil. For example, leaf litter on the soil does not change the reflectance properties nor does water content. Soil reflectances in SiB2.5 are used in the calculation of surface albedo and consequently available energy. Sensitivity to these parameters is exhibited year round with the LE,H criterion and in particular in May, November and December for both criterion. In the months with extreme sensitivity, the reflectance of the soils is pushed into the lower part of their range, decreasing albedo and therefore increasing the amount of energy entering the ground and decreasing the amount of energy scattered back through the canopy (Figure 3.11). The opposite effect occurred in April.

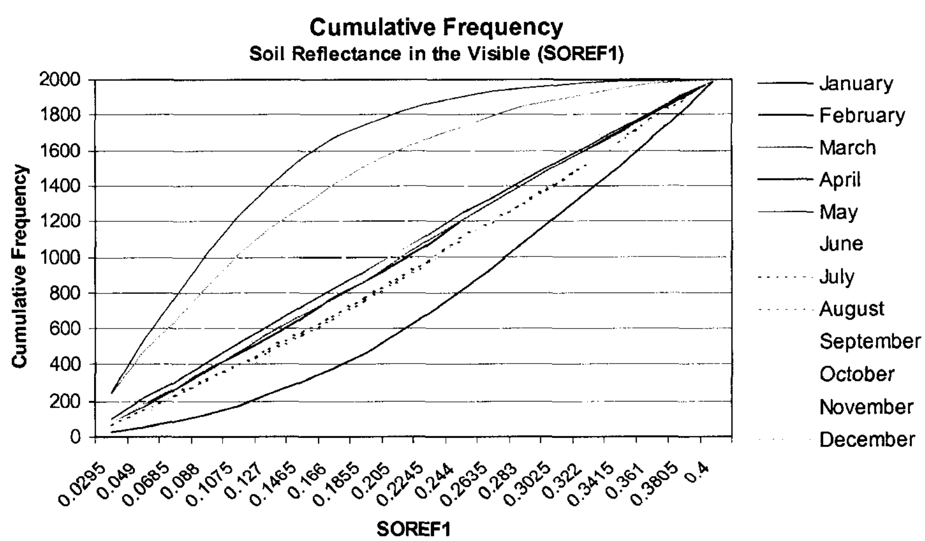


Figure 3.11 Cumulative frequency diagram of the top 2000 fitted values for the parameter *SOREF1,* the LE,H criterion. May, October, November and December exhibit strong sensitivity and adjustment towards lower reflectance. April exhibits strong sensitivity as well but adjustment toward higher reflectance.

SiB2.5 was also found to be sensitive to the specification of the leaf angle distribution (*CHILL*), though much less so than to the soil reflectances. CHILL is used to calculate *GMUDMU* (Appendix 2), one of the time varying parameters. *GMUDMU* is used in the calculation of the distribution factor for radiation and other properties within the canopy. SiB2.5 appears to be most sensitive to *CHILL* in the winter and in the shoulder seasons. In this experiment, *CHILL* was varied from –0.5 to 0.5, where negative numbers indicate more vertically oriented leaves and positive numbers indicate more horizontally oriented leaves. In the shoulder seasons and winter, preferred values of *CHILL* tend toward negative values (Figure 3.12), or a more vertical canopy, perhaps allowing more light to reach the surface.

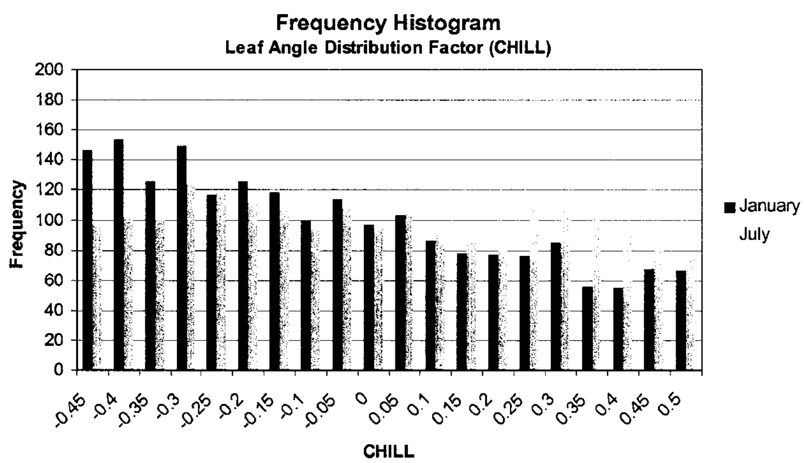


Figure 3.12. Frequency histogram of the top 2000 fitted values for the parameter CHILL, the LE,H criterion. January (Black) and July (grey) are shown as examples for Winter and Summer respectively.

VCOVER is used off-line in the calculation leaf area index and f_{PAR} (Appendix 2) and in-line to calculate the surface cover weighted proportions for radiative transfer (visible, near-infrared and thermal). *VCOVER* shows sensitivity virtually year round without any extremes. The preferred values of *VCOVER* change both through the seasons and between the two criteria, but the changes are not very large.

3.4.6 *Soil moisture availability parameters*

The most sensitive parameter in this category is root depth (*ROOTD*). *ROOTD* is used in the calculation of soil layer thickness and therefore has an impact on heat transfer through the soil as well as soil respiration. *ROOTD* maintained a near constant level of sensitivity throughout the year and, in general, shallower depths were favored. *SODEP*, which is used to calculate the layer depths for soil moisture, was only sporadically sensitive (April and the winter months). In those time periods, deeper soil depths were favored.

The model was also quite sensitive to the parameter *BEE*, the soil wetness exponent. *BEE* is an exponential parameter that is used along with saturated soil water potential (*PHSAT*) and soil wetness to determine the unsaturated soil water potential (Clapp and Hornberger 1978). The water potential of any soil is a measure of how much work must be done to remove water from the soil. Less negative values of *BEE* are likely to be found with sandier soils and indicate easier movement. More negative values of *BEE* are likely to be found with more clay-like soils and indicate more difficult movement of water through the soil.

In this experiment, SiB2.5 is sensitive to the parameter *BEE* in January, April and November. The preferred range for *BEE* in these months is shifted toward the lower part of the range of *BEE*, those values that reflect sandier soils and easier water movement. April and November are months where the soils went through freeze/thaw cycles at WLEF. Water in the soil has a higher heat capacity than soil does, so wet soil takes longer to heat up and longer to cool down than dry soil. The effect of water on the conductivity of heat in the soil is more non-linear, but does serve to increase the conductivity of heat. It is possible, then, that this combination of parameter values, *BEE* plus the changing soil and root depths, reflects an adjustment for moisture status and the regulation of soil temperatures in these months.

3.4.7 *Photosynthesis parameters*

The photosynthesis parameters were quite well behaved. The parameter *HHTI*, the photosynthesis high temperature inhibition ½ point, displayed a very consistent normal distribution with a peak that shifted between 300 and 310 degrees Kelvin (Figure 3.13a). The stability of the distribution through time suggests that there is a relatively fixed optimal value for this parameter. The photosynthesis low temperature inhibition ½ point (*HLTI*), was also consistent in that the favored values were in the upper half of the range. However, in summer the distribution was pushed even more toward the upper end of the distribution of *HLTI* (towards 290 °K) (Figure 3.13b). One characteristic of SiB2.5 that has been noted in the past (Baker *et al.* 2003) is too rapid resumption of CO_2 uptake in the morning and too slow cessation of uptake in the evening during the summer months. This has been attributed to problems with radiative transfer in the canopy at low sun angles. It is possible that increasing the low temperature inhibition ½ point would address this issue from a thermodynamic rather than light availability perspective.

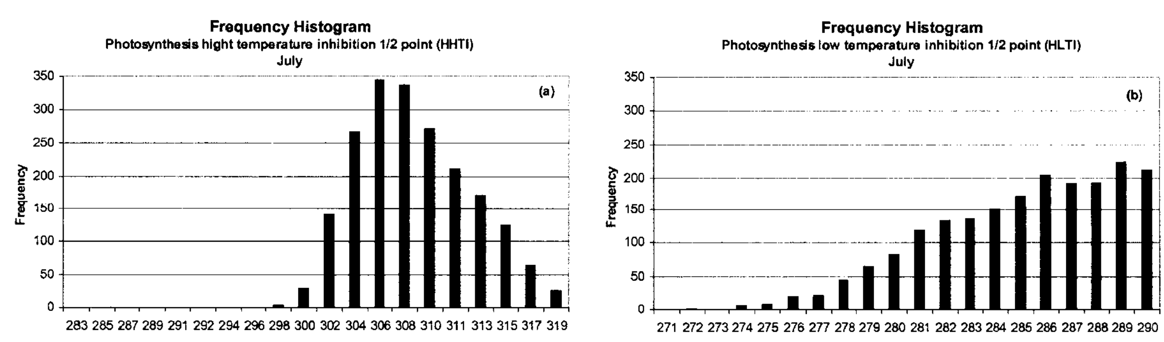


Figure 3.13. Frequency histograms of the top 2000 fitted values of the parameters *HHTI* (a) and *HLTI* (b) for the month of July and the NEE, LE, H criterion.

Three parameters that control the rates and efficiency of photosynthesis: the rubisco velocity of a leaf in full sun (*VMAX*), the slope of the Ball-Berry photosynthesis-conductance relationship (*GRADM*) and the minimum stomatal conductance (*BINTER*), show the same broad patterns. They are all highly sensitive during the growing season but also show a peak in sensitivity in April. This peak in April appears to be related to previously described sensitivity peaks in April where anomalous parameter distributions were identified. The parameter ranges indicated for April suggest that photosynthesis is being curtailed. Remember that the remote sensing parameters from the top parameter sets produced more leaf area and higher f_{PAR} in April, despite the fact that the only leaves on the vegetation would have been on evergreens as the growing season had not yet begun. This result suggests that the changes in the remote sensing parameters in April were related to energy balance rather than net ecosystem exchange.

During the growing season, the parameters *VMAX* and *GRADM* undergo shifts in their ranges that appear to be ecologically meaningful. At the beginning and the end of the growing season, best-performing parameter sets have lower parameter values (indicating reduced photosynthetic capacity and reduced conductance) for these two parameters than during the height of the growing season. Results from a recent model data comparison with SiB2 in an agricultural-prairie system indicate that photosynthetic capacity was overestimated by SiB2 at the beginning and end of the growing season and that both young and senescing leaves might have a more reduced photosynthetic capacity than high-season leaves (Hanan *et al.* in prep). These results suggest the same (Figure 3.14a.b).

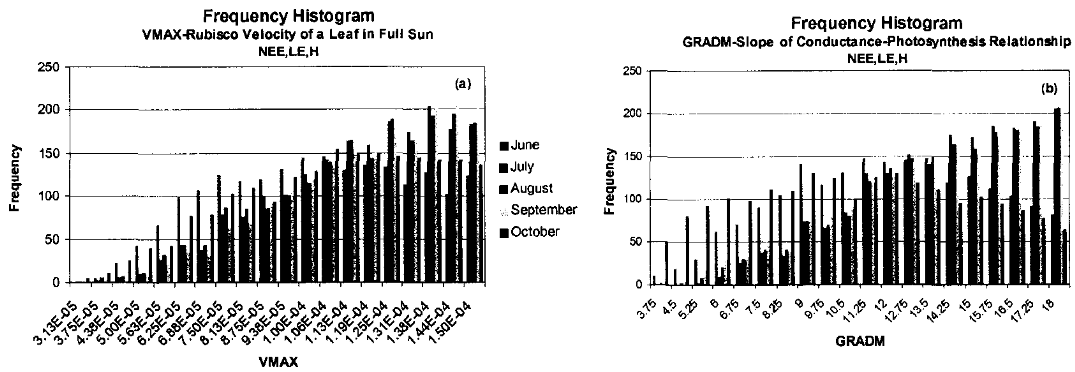


Figure 3.14. Frequency histograms of the top 2000 fitted values of the parameters *VMAX* (a) and *GRADM* (b) for the growing season and the NEE, LE, H criterion. The red and blue bars represent the beginning and end of the growing season (June and October respectively), the gray bars represent the height of the growing season (July, August, September). Notice the two distinctly different distributions.

The quantum efficiency (*EFFCON*) parameter also shows sensitivity during the growing season but, unlike the three parameters mentioned above, has a sensitivity peak in May. This peak appears to be related to the premature greening up in May and associated overestimation of NEE since this sensitivity peak is only seen with the NEE,LE,H criterion. Further, the peak represents an adjustment of *EFFCON* to lower quantum efficiencies, unlike during the growing season where the best runs have quantum efficiencies in the top half of the parameter range. Similarly, the parameter *VMAX* in May exhibits a strongly asymmetric distribution favoring the lowest part of the parameter range and thus lower rates of conductance and assimilation.

3.4.8 *Autotrophic respiration parameters*

As in the set of photosynthesis parameters, there were some well-behaved parameters in the autotrophic respiration set. *TRDM*, the autotrophic respiration inhibition ½ point temperature, was sensitive across all months and had a normal distribution centered consistently on 321° K (Figure 3.15a). Interestingly, this temperature is about seven degrees lower than the value assigned for a typical WLEF simulation. The stability of the distribution through time suggests that there is a stable optimal value for this parameter, much like *HHTI.* The optimum temperature for autotrophic respiration (primarily leaf respiration in SiB2.5), *TROP*, which is used in the Q_{10} equation for autotrophic respiration, was sensitive to a variable degree. For the LE,H criterion, the sensitivity was confined to the growing season, for the NEE, LE, H criterion the sensitivity extended all year. During the growing season, the distribution of *TROP* centered on 298°K, the default value for WLEF. Outside of the growing season however, the upper half of the distribution, or higher optimum temperatures for autotrophic respiration, were favored. Since these are generally colder months, these results would indicate adjustment for less leaf respiration in dormant months (Figure 3.15b).

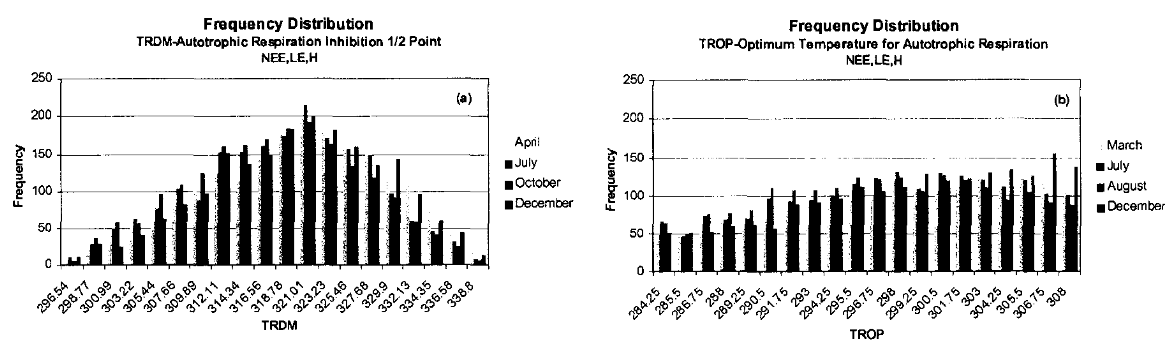


Figure 3.15. Frequency histograms of the top 2000 fitted values of the parameters *TRDM* and *TROP* **a)** *TRDM*, the autotrophic respiration inhibition ½ point. Notice how the normal distribution remains similar across many months with a peak around 321°K. **b)** Frequency histogram of the parameter *TROP*, the optimum temperature for autotrophic respiration. Notice that during the growing season, the distribution is nearly normal around 298°K whereas in the dormant season the distribution is pushed towards higher temperatures.

Like *TROP*, *RESPCP* was sensitive during the growing season for the LE,H criterion, but virtually all year for the NEE,LE,H criterion. *RESPCP* is the leaf respiration fraction of *VMAX.* In general, the favored distribution for *RESPCP* was the lower half of its range, also limiting leaf respiration. The limitation on leaf respiration could potentially be offsetting too high heterotrophic respiration at these times

3.4.9 *Heterotrophic respiration parameters*

An unanticipated outcome was that the modeled results were not sensitive to the parameters that regulate heterotrophic respiration. Heterotrophic respiration in SiB2.5 is a function of three things, the net annual assimilation of carbon, soil moisture and soil temperature. In SiB2.5, annual assimilation of carbon is entirely offset by respiratory losses, leaving a net zero accumulation of carbon in the ecosystem. The respiratory losses are parameterized diurnally and seasonally by partitioning the annual assimilation between the six soil layers and the surface and regulating the respiration rate through temperature and moisture (Raich and Nadelhoffer 1989; Denning *et al.* 1996a). The respiration rate is controlled by a simple exponential relationship for temperature (commonly referred to as Q_{10}) and by three parameters related to soil texture (*WOPT* – the optimal percent soil saturation for respiration, *WSAT* – the heterotrophic respiration rate at saturation and *ZM* – the skewness exponent of the soil moisture stress factor), which partition respiration according to available soil moisture. ZM, the skewness exponent, is a statistical parameter that adjusts the respiration rate based on clay fraction. In this experiment, we have varied the values of *WOPT, WSAT* and *ZM* but not Q_{10}, which remained constant at 2.4. In SiB2.5, Q_{10} is constant across soil textures and biomes and therefore not considered an input parameter, whereas *WOPT*, *WSAT* and *ZM* vary. One could argue that Q_{10} should also have been varied in this experiment, but it is likely that this would not have changed the result much. Figure 3.16 shows the results of simulations utilizing the optimal parameter set for each month but varying the Q_{10}. In each case the simulations were conducted as described previously but the Q_{10} was changed. Figure 3.16 shows that for these sets of parameters, varying the Q_{10} did not greatly affect the RMSD error. Figure 3.17 shows the actual fluxes used to calculate the RMSD errors in Figure 3.16 for March and July. The variability in the observations overwhelms the comparatively small differences the variations in Q_{10} generate.

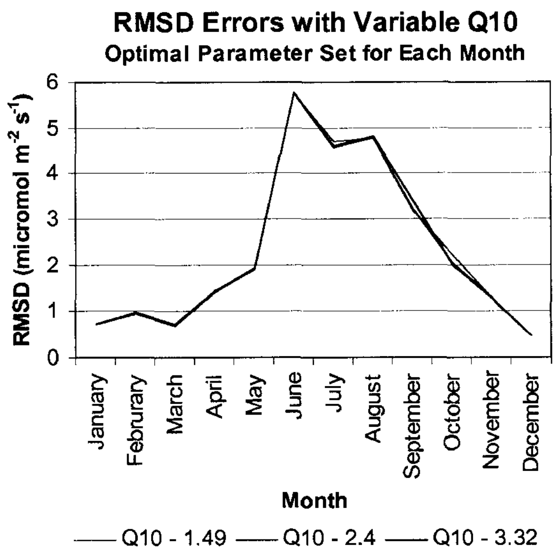


Figure 3.16 RMSD errors for simulations using the optimal parameter set for each month but varying the soil Q_{10}.

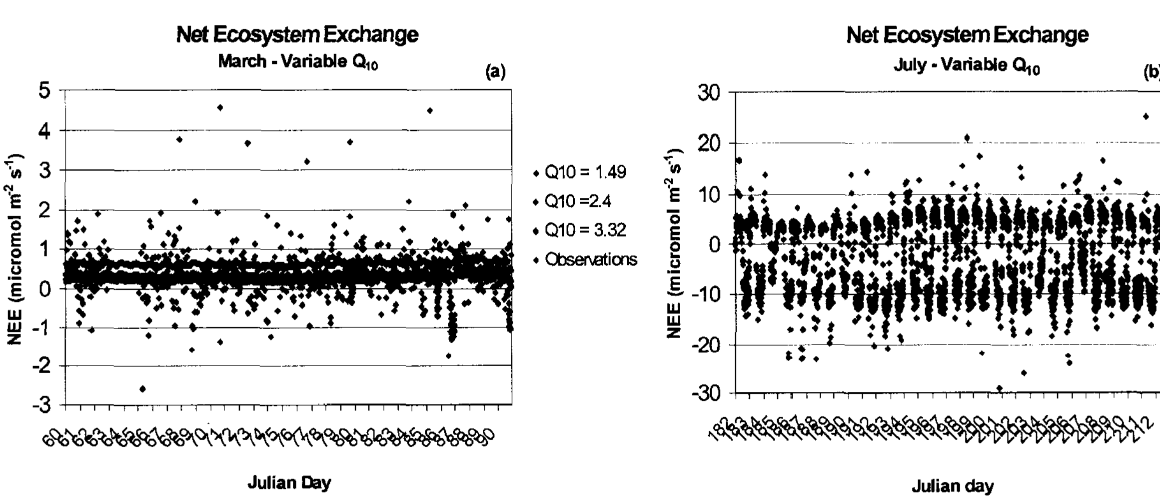


Figure 3.17a,b Simulated fluxes using the optimal parameter set for each month but varying the soil Q_{10}. Observations are shown in Grey.

The insensitivity of the results to the heterotrophic respiration parameters that were varied in this analysis is likely due to: 1) the requirement that annual net assimilation and respiration balance 2) the spin-up of the model and 3) compensation by other parameters. Since the calculation of respiration is based on the net annual assimilation, it follows that the model has to first be run, the total net annual assimilation calculated, and then run again with this valued being partitioned into heterotrophic respiration. The partitioning is done through the terms respfactor and soilscale. Respfactor is the basic increment of respiration that the temperature and moisture effects are multiplied by. Soilscale is the integrated effect of temperature and moisture that is summed through the year and used to adjust the respiration increment in respfactor. In this analysis, all parameter sets were spun up for six years to equilibrium, with respfactor and the initial conditions being updated each year, and then run a final time. With a spin-up, adjustments to respfactor for the different parameter sets are effectively made during the spin-up each year such that if soil moisture is low or the soil temperature is low, respfactor is adjusted for this via soilscale. The result is that the model is not sensitive to the heterotrophic respiration parameters.

A sensitivity analysis conducted in the identical fashion but made without any model spin-up demonstrates this idea. Figure 3.18a,b shows the heterotrophic respiration parameters for both the spun up and not spun up version of the sensitivity analysis. Keep in mind that in the version that was not spun up, the annual heterotrophic respiration violates the assumption that it must balance with net annual assimilation and the soil moisture and temperature fields were not equilibrated. When the model has not been spun up it shows much greater sensitivity to the heterotrophic respiration parameters that were varied in this analysis (Figure 3.18b) than it does when it has been spun up (Figure 3.18a).

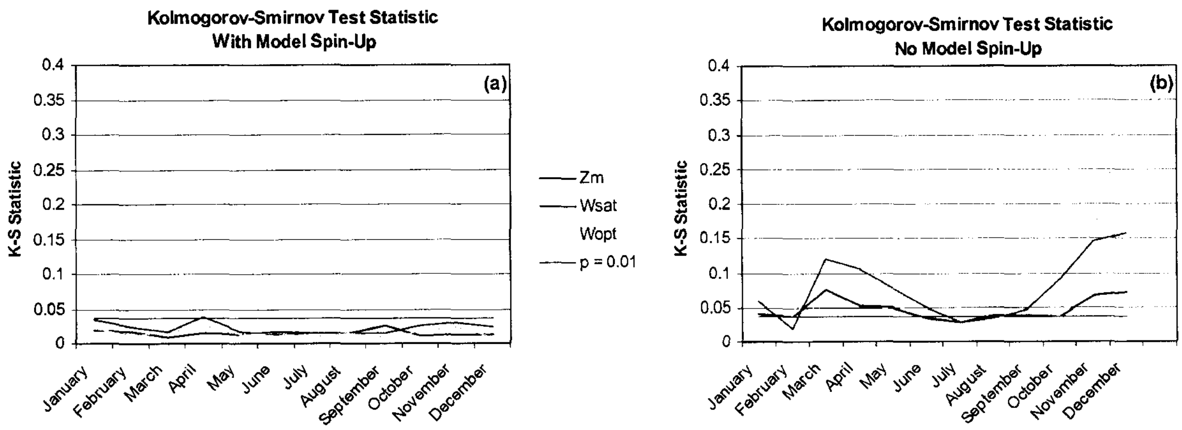


Figure 3.18. Kolmogorov-Smirnov test statistics for parameters *ZM*, *WSAT* and *WOPT* for all twelve months for the NEE,LE,H criterion. **a)** with a six year spin-up of the model **b)** without any spin-up.

3.4.10 *A note on the month of April*

The month of April appears to be difficult for SiB2.5 to model, even with large freedom to vary model parameters. Results in April show great sensitivity to a wide variety of parameters that, taken as a whole, are difficult to interpret. The optimal parameter sets resulted in high f_{PAR}, little photosynthesis and highly reflective, deep, sandy soils. At the WLEF site, the buds on the deciduous trees had not yet burst, snow was melting and the soil thawed midway through the month. The strange parameter sets and extreme sensitivity point to problems with the model structure or logic. One known problem with SIB2.5 is that when snow first becomes patchy, too much radiation is absorbed by the small amount of visible ground, producing implausible temperature spikes. In the standard WLEF run, the first sign of this appears at the very end of March (reaching 37°C in the standard run) and carries on into the first week of April (persistent 25°C, then settling to 17°C by the end of the month). Since the driver data is the same across all simulations (e.g. same precipitation, same air temperatures) it is possible that the optimal parameter sets for April are compensating for this problem with the unrealistic adjustments in f_{PAR} and soil reflectivity.

3.4.11 *Parameter Rankings*

An advantage of the Kolmogorov-Smirnov test statistic is that it can be used to rank the parameters (Tables 3.6 and 3.7). The higher the Kolmogorov-Smirnov test statistic, the more statistically significant it is (Hornberger and Spear 1981) and the more sensitive the model is to the specification of that parameter. Many of the top influential parameters are shared by both criteria and are therefore perhaps the most important to define well. Considering only the ten most influential parameters in each month, three parameters stand out as present in this group in each month and for both criteria, these are: *ROOTD*, *Z1* and *HHTI*. Other parameters in this most influential goup that the two criteria share at least half the time (e.g. 6 months or more) are: *ZC*, *TRDM*, *SOREF2*, *Z2*. It is interesting to note that *HHTI* and *TRDM*, which are both photosynthesis parameters, are highly influential for both criteria, indicating that the addition of CO_2 is important in constraining the energy as well as the carbon fluxes. Other parameters are influential to both criteria, but for shorter periods of time such as the growing season or shoulder seasons. Depending on the modeling objectives (long or short term, carbon or energy, predictive or diagnostic) it might be important to improve the accuracy of some parameters over others.

Table 3.6. Model parameters from the top 2000 parameter sets ranked for each month by the Kolmogorov-Smirnov coefficient for the LE, H multiobjective criteria. Parameters are ranked in descending order of influence. Parameters shaded in orange are significant at the p=0.01 level.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All
z2	z2	rootd	nd98	soref2	hhti	hhti	hhti	hhti	zc	z2	z2	hhti
zc	zc	nd98	rootd	sorefl	hlti	gradm	gradm	rootd	z2	zc	zc	hlti
zl	zl	z2	gradm	zl	rootd	hlti	hlti	gradm	nd98	zl	zl	rootd
rootd	rootd	zc	sorefl	zc	zl	rootd	rootd	zl	zl	soref2	soref2	gradm
hhti	hhti	zl	ltmax	rootd	zc	vmax0	zl	vmax0	rootd	chil	sorefl	zl
trdm	soref2	stem	zl	z2	gradm	zl	vmax0	hlti	chil	sorefl	sodep	zc
chil	trdm	nd02	soref2	hhti	soref2	respcp	zc	zc	soref2	sodep	rootd	z2
soref2	nd98	chil	zc	lwidth	trdm	zc	respcp	vcover	hhti	bee	hhti	vmax0
vcover	stem	hhti	vmax0	trdm	vcover	trdm	z2	nd98	trdm	rootd	trdm	soref2
bee	nd02	trdm	hhti	vcover	binter	vcover	trdm	z2	ltmax	hhti	bee	trdm
nd98	chil	soref2	trdm	gradm	respcp	soref2	vcover	respcp	lwidth	trdm	nd98	respcp
poros	sodep	ltmax	bee	ref21	vmax0	binter	lwidth	lwidth	sodep	vcover	poros	vcover
sodep	ltmax	vcover	chil	chil	nd98	effcon	nd98	trdm	sorefl	satco	satco	binter
lwidth	lwidth	tran21	stem	hlti	trop	trop	soref2	soref2	bee	lwidth	tran21	trop
satco	tran21	ref22	nd02	nd98	z2	z2	binter	binter	ref21	gradm	stem	lwidth
tran21	bee	ref21	binter	tran21	poros	sorefl	effcon	ref21	gradm	llength	ltmax	effcon
refl2	ref21	sodep	sodep	bee	effcon	ref21	trop	effcon	tran21	poros	ref21	chil
llength	llength	tran22	ref21	vmax0	lwidth	btheta	ref21	btheta	binter	phsat	trop	nd98
effcon	slti	hlti	respcp	stem	shti	atheta	sorefl	sorefl	satco	tran21	tran11	stem
phsat	shti	btheta	satco	satco	sodep	tran21	ltmax	ltmax	poros	vmax0	refl1	sodep
binter	binter	bee	tran21	binter	ltmax	ph	ph	tran21	refl1	trop	slope	sorefl
ref22	respcp	gradm	vcover	refl1	satco	lwidth	btheta	refl1	nd02	nd98	lwidth	ltmax
sorefl	refl2	tran11	trop	sodep	sorefl	refl1	tran21	ph	ref22	ref21	zm	nd02
ltmax	ref22	poros	hlti	effcon	atheta	bee	atheta	atheta	trop	wopt	respcp	satco
respcp	sorefl	refl2	lwidth	trop	ph	chil	chil	stem	llength	nd02	binter	ph
tran12	satco	llength	llength	atheta	chil	poros	stem	trop	tran22	slti	btheta	atheta
stem	poros	lwidth	poros	respcp	btheta	stem	refl1	bee	btheta	refl2	chil	poros
slope	wopt	sorefl	z2	shti	tran22	ref22	bee	sodep	refl2	binter	vcover	tran11
trop	tran22	tran12	ph	slti	slti	sodep	poros	chil	vcover	atheta	ref22	ref22
btheta	effcon	effcon	phsat	btheta	tran21	trda	nd02	shti	respcp	ph	effcon	bee
refl1	atheta	vmax0	atheta	tran11	phsat	ltmax	shti	nd02	stem	ltmax	hlti	btheta
wopt	trop	respcp	tran11	poros	slope	satco	satco	satco	slope	slope	tran22	refl2
gradm	hlti	wsat	wsat	zm	bee	tran11	refl2	ref22	shti	respcp	phsat	tran21
nd02	slope	slope	effcon	phsat	llength	wsat	trda	llength	hlti	wsat	vmax0	tran12
wsat	ph	refl1	shti	slope	wsat	slti	llength	tran11	slti	trda	nd02	tran22
vmax0	tran12	satco	wopt	trda	tran11	refl2	tran12	poros	wopt	stem	wopt	phsat
trda	refl1	atheta	refl1	tran22	ref21	nd98	ref22	wsat	zm	tran22	gradm	llength
shti	gradm	phsat	trda	llength	nd02	phsat	sodep	tran22	atheta	tran12	shti	wsat
zm	btheta	wopt	slope	refl2	ref22	slope	slope	slti	tran12	shti	slti	zm
slti	vmax0	zm	btheta	ph	tran12	wopt	wopt	trda	vmax0	effcon	ph	trda
tran22	phsat	binter	zm	tran12	stem	nd02	phsat	phsat	phsat	tran11	tran12	wopt
hlti	wsat	shti	ref22	wopt	wopt	shti	zm	tran12	ph	refl1	refl2	slti
atheta	vcover	slti	slti	ltmax	refl2	tran12	slti	slope	wsat	hlti	atheta	ref21
tran11	trda	ph	tran12	nd02	trda	tran22	wsat	wopt	effcon	zm	trda	refl1
ph	zm	trop	tran22	wsat	refl1	llength	tran11	zm	tran11	ref22	llength	slope
ref21	tran11	trda	refl2	ref22	zm	zm	tran22	refl2	trda	btheta	wsat	shti

Table 3.7. Model parameters from the top 2000 parameter sets ranked for each month by the Kolmogorov-Smirnov coefficient for the LE, H, NEE multiobjective criteria. Parameters are ranked in descending order of influence. Parameters shaded in orange are significant at the p=0.01 level.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All
z2	z2	rootd	vmax0	soref2	hhti	hhti	hhti	hhti	hhti	zc	hlti	hhti
zc	zc	hhti	nd98	vmax0	hlti	hlti	hlti	vmax0	z1	z2	hhti	hlti
z1	hhti	stem	rootd	sorefl	rootd	gradm	gradm	rootd	zc	z1	z1	rootd
rootd	z1	hlti	gradm	rootd	z1	vmax0	vmax0	gradm	z2	hhti	zc	vmax0
hhti	rootd	zc	hhti	z1	vmax0	rootd	rootd	hlti	rootd	soref2	soref2	z1
trdm	hlti	z2	z1	nd98	zc	z1	z1	z1	nd98	chil	z2	gradm
respcp	respcp	z1	sorefl	hhti	gradm	respcp	zc	zc	vmax0	sodep	rootd	zc
soref2	trdm	nd98	trop	zc	soref2	zc	respcp	vcover	respcp	trdm	sorefl	trdm
chil	trop	trdm	hlti	trdm	trdm	trdm	trdm	trdm	soref2	rootd	sodep	soref2
vcover	nd98	vmax0	trdm	hlti	nd98	soref2	soref2	respcp	trdm	respcp	trdm	respcp
trop	soref2	trop	zc	effcon	binter	effcon	z2	soref2	chil	sorefl	trop	trop
nd98	stem	soref2	stem	slti	respcp	binter	lwidth	z2	hlti	bee	vmax0	binter
hlti	chil	chil	ltmax	gradm	trop	trop	vcover	effcon	slti	trop	respcp	z2
bee	effcon	respcp	soref2	z2	vcover	vcover	effcon	lwidth	lwidth	vcover	stem	effcon
sodep	ltmax	nd02	binter	ref21	effcon	btheta	binter	binter	bee	hlti	nd98	vcover
poros	btheta	ltmax	bee	btheta	ph	ph	trop	btheta	sodep	slti	bee	chil
vmax0	nd02	vcover	sodep	trop	poros	atheta	btheta	ref21	ltmax	satco	vcover	nd98
sorefl	bee	gradm	ref21	ltmax	sodep	ref21	nd98	ph	vcover	lwidth	refl1	ph
btheta	lwidth	tran21	chil	shti	tran22	sorefl	ph	nd98	effcon	nd98	ref21	lwidth
wsat	tran21	ref22	ph	respcp	shti	nd98	atheta	sorefl	trop	vmax0	ltmax	sodep
lwidth	tran11	sodep	nd02	atheta	btheta	tran21	sorefl	trop	binter	llength	slti	atheta
ltmax	ref22	ref21	lwidth	vcover	chil	z2	ref21	atheta	shti	gradm	chil	poros
effcon	atheta	effcon	vcover	lwidth	sorefl	refl1	chil	refl1	satco	phsat	poros	satco
llength	gradm	refl2	btheta	stem	atheta	bee	ltmax	tran21	gradm	wsat	ph	tran21
stem	vmax0	ph	tran21	tran21	ltmax	lwidth	tran21	stem	stem	effcon	effcon	btheta
tran11	wsat	tran11	satco	satco	satco	poros	poros	bee	poros	binter	slope	slti
gradm	ref21	tran12	shti	ph	slope	refl2	refl1	ltmax	ph	poros	tran21	refl2
refl1	satco	refl1	wsat	bee	lwidth	chil	stem	zm	btheta	wopt	wsat	stem
tran12	slope	binter	tran11	chil	refl2	ltmax	bee	chil	ref21	shti	binter	ref22
binter	binter	trda	slti	llength	tran21	ref22	phsat	sodep	tran21	ref21	satco	sorefl
phsat	vcover	sorefl	atheta	refl1	bee	stem	trda	tran11	refl2	nd02	gradm	ltmax
satco	tran12	slope	poros	nd02	trda	trda	refl2	poros	ref22	tran21	trda	ref21
wopt	phsat	bee	z2	sodep	llength	satco	shti	tran12	sorefl	ltmax	shti	tran22
zm	zm	tran22	wopt	slope	refl1	tran22	tran11	shti	wsat	tran12	phsat	nd02
slti	sodep	wsat	refl1	tran12	zm	sodep	zm	wsat	atheta	slope	lwidth	slope
tran22	slti	llength	trda	wsat	ref21	llength	wsat	slti	wopt	tran11	tran11	trda
slope	llength	lwidth	phsat	poros	stem	wsat	wopt	slope	trda	tran22	nd02	bee
ref21	refl1	poros	effcon	tran11	phsat	shti	sodep	refl2	llength	trda	llength	wsat
shti	wopt	atheta	llength	binter	wopt	zm	tran12	satco	tran11	atheta	ref22	wopt
ref22	sorefl	wopt	slope	tran22	z2	phsat	llength	nd02	tran22	refl2	atheta	llength
tran21	trda	btheta	respcp	wopt	nd02	tran11	nd02	tran22	slope	stem	zm	refl1
ph	ph	phsat	zm	zm	slti	nd02	satco	ref22	phsat	zm	tran12	phsat
refl2	refl2	satco	tran12	ref22	wsat	tran12	slope	trda	zm	refl1	btheta	shti
trda	tran22	shti	refl2	refl2	tran12	slope	slti	wopt	nd02	ref22	refl2	tran11
atheta	shti	zm	ref22	trda	ref22	wopt	tran22	llength	tran12	btheta	tran22	tran12
nd02	poros	slti	tran22	phsat	tran11	slti	ref22	phsat	refl1	ph	wopt	zm

3.5 Discussion

Analyses such as the one presented here are very sensitive to the observations against which the model is being tested. The eddy covariance data used in this dissertation are not raw data; they are corrected for a variety of phenomena. These corrections include such things as: 1) coordinate rotation, 2) time lag to the sensor, 3) frequency response and 4) storage terms (Berger *et al.* 2001; Davis *et al.* 2003). When changes to these corrections occur, the observations change and therefore the results of any analysis that uses them. In the course of this work, the investigators providing the observations changed their analyzed fluxes numerous times producing sometimes subtle and sometimes quite obvious differences. Similarly, simulation results are dependent to a large extent on the observations used to drive the model (i.e. driver data). Model driver data must be continuous; observations made in the field never are. Instrument breakdowns, dropped data and environmental issues (freezing, flooding, fire) all cause gaps in data that need to be filled. Even though the filling of driver data is done as rigorously as possible (see Chapter 2 for a description of how the WLEF driver data were filled), large gaps remain a very difficult problem.

The results of model sensitivity analyses such as this one are also dependent on both the site that is being modeled and the weather conditions during the simulation, because the strengths and weaknesses of any model vary with both of these conditions. Drought or flood conditions that produce anomalous environmental conditions might not be a good predictor of how well a model does at characterizing a site in more average years or how sensitive a model is generally to its parameterization. On the other hand, an anomalous year might help to fine tune parameterizations and/or reveal other deficiencies in model structure and logic not previously identified. Further, the sensitivity a model displays in one biogeographical location will probably not translate directly to another.

Despite the many limitations, which oftentimes cannot be avoided, analyses such as these can be very powerful. This sensitivity analysis identified many parameters whose specifications had a significant influence on modeled results. Patterns in the sensitivity of the parameterization were related to many things: the ecology of the system, the specification of input data and how they were used, and weakness in the model structure itself. As a result, it is difficult to determine in many cases whether better specification of a parameter or better representation of a process is more essential to improving model results. Only in a few cases was this clear. For example, the unusual optimal parameter sets for April made it evident that there were environmental conditions and processes that the model was struggling to simulate well in that month. In contrast, the parameters *HHTI* and *TRDM* were remarkably consistent and simple adjustment of the parameter values themselves is all that seems necessary for improvement. For the rest of the cases, the interpretation is quite complex and some reciprocation between parameter specification and process representation is probably necessary.

Unlike many sensitivity analyses, which have focused on time invariant parameterizations, the one presented here investigated how sensitivities change through time. In SiB2.5 a combination of time varying and time invariant parameters are utilized. However, the time varying parameters represent only those that relate to vegetation cover via the normalized difference vegetation index. The results presented here show that model sensitivity clearly varies through time for a much larger set of parameters. In some cases, as in the month of April, the sensitivity may not be ecologically meaningful but may serve to improve model results by pointing out model structural deficiencies. In other cases, the results were both affirming and meaningful. One would expect parameter dominance for different land surface processes in different seasons, and this was evident in the parameter rankings that showed parameters related to photosynthesis more influential in the growing season and parameters related to energy balance more influential at other times of the year. The variability of the top parameter sets within those seasons provided additional useful information. In particular, the within season variability of the top parameter sets for the photosynthesis parameters *VMAX*, *GRADM*, *HLTI* drew attention to biophysical characteristics that behave quite differently than expected. It has generally been assumed that properties of photosynthesizing vegetation such as *VMAX*, *GRADM* and *HLTI* don't vary with the age or stage of the vegetation; the results presented here suggest an interesting topic for further research. Similarly, it has long been felt that the mapping of soil reflectance by biome was a weak point in the parameterization of SiB. The sensitivity to soil reflectance that the model exhibits reinforces this sentiment.

Given the number of parameters investigated, it was not altogether surprising to find that SiB2.5 exhibited signs of equifinality. At any given time, half of the parameters were found to be non-influential. It is encouraging that if reasonable values are known, we may not need to worry too greatly about some parameters. However, the influential and non-influential parameters varied by month and would also most likely vary at shorter and longer time scales and for different vegetation types or different environmental conditions. Further, non-influential does not mean that parameters are unimportant. If the non-influential parameters contribute to the mechanistic underpinnings of the model, even if only in a minor way, they are still important, particularly when natural conditions change in unpredictable ways.

The addition of net ecosystem exchange (NEE) does appear to help constrain the model parameterization and thus the prediction of energy and water fluxes. In the single point version of SiB2.5, energy is conserved at that point. If either the heat or water flux is over or under predicted, the remainder of the available energy is committed to the other flux. In the formulation of SiB2.5, NEE directly influences latent heat flux and indirectly influences sensible heat flux. The high ranking of photosynthesis parameters for the LE,H criterion in the growing season (Table 3.6) are evidence of the additional, positive constraints NEE places on the fluxes of latent and sensible heat in SiB2.5.

The results of this sensitivity analysis also emphasize the risk involved in undirected model optimization. Taken one step further, the results presented here could be used to optimize SiB2.5. Yet, after careful analysis it is clear that optimization of all the parameters would not be entirely biologically correct. Parameter interactions are very non-linear and complex and besides compensating for one another, they compensate for imperfect driver data and model structural problems. An analysis such as this can help pare down the parameters for which optimization is reasonable, which is important because a good fit to data for the wrong reasons would strip a model of its predictive power.

In the next chapter, the results of this sensitivity analysis are used to examine the uncertainty in simulations of surface fluxes that are a direct result of the parameterization. With knowledge of model sensitivity, we should be better able to interpret the origin of the uncertainty and in particular how and why uncertainty varies in time.

CHAPTER 4

UNCERTAINTY ANALYSIS

4.1 Introduction

Scientists use models of natural systems to make decisions about the accuracy of hypotheses, to recommend a course of action or simply to learn about how the system works. Uncertainty is an intrinsic part of this decision making process. Models can never reflect absolute reality; therefore interpretation of modeled results must incorporate some level of uncertainty, some acknowledgment of the probable versus the unknown.

In a recent special section of the journal *Ecology,* Brewer and Gross (Brewer and Gross 2003) argue that ecologists need to be trained to "think with uncertainty in mind". They suggest that as scientists who are expected to play a large role in determining the consequences of global climate change, ecologists must be able to understand how the various kinds of uncertainty impact our models and forecasts and ultimately the decisions that will be made from them. There is also a growing body of literature in ecology that suggests that actually incorporating uncertainty into deterministic models by allowing for some stochastic or unknown effects produces more realistic and robust predictions (Clark 2003; Clark *et al.* 2003).

However, uncertainty is a difficult and challenging concept. The stochastic processes that play such a large part in the functioning of natural systems are difficult to quantify because random occurrences can cause cascading effects that might not be fully imaginable. In deterministic models, space, time and rate limits are set to efficiently characterize processes; this curtails our ability to incorporate uncertainty and paradoxically engenders uncertainty (Kremer, 83). Further, the motivation for ecologists to add uncertainty to already complex models and interpretations is often just not there. As Fedra et al. (Fedra *et al.* 1981) eloquently stated, "what choice would any of us make between an honest but depressingly ambiguous prediction and a faithfully promoted, complex and seemingly accurate prediction?"

Difficulties aside, though, it is becoming ever clearer that we must incorporate uncertainty into our modeling systems. In the same special section of *Ecology*, Pielke Jr. and Conant (Pielke Jr. and Conant 2003) describe an example where incorporating model uncertainty into the decision making process would have substantially improved the outcome. In their example on flood prediction, if the calculable uncertainty had been incorporated into the river cresting height estimates, precautions for flooding might have been more substantial and thus more damage averted. They also show how uncertainty is successfully used in other disciplines, such as atmospheric science. The example they give is weather prediction, where uncertainty is reported with the daily forecast as well as with extreme events. In weather prediction, uncertainty is expected and therefore the interpretation of that uncertainty is accessible to those who use these forecasts (Pielke Jr. and Conant 2003).

In carbon cycle modeling, the impact of poor uncertainty estimation and interpretation might not be as immediate as in flood forecasting or weather prediction, but that does not make it any less pressing that we develop full comprehension of it. Two of the most high profile research topics in science today are: 1) what are the possible environmental responses to increased CO_2 and the likely trajectories of change (Houghton *et al.* 2001; McCarthy *et al.* 2001) and 2) how and where are the land surface and oceans currently mitigating the increases in efflux of CO_2 into the atmosphere (Schimel 1995; Fan *et al.* 1998; Tans and White 1998; Schimel *et al.* 2000; Valentini *et al.* 2000; Pacala *et al.* 2001; Gurney *et al.* 2002). As a scientific community, we are beginning to understand the processes involved and the interacting effects, but the uncertainty bounds on both the expected increase of CO_2 in the atmosphere and on the amount of carbon the oceans and land surface will be able to absorb are still large and controversial (Kheshgi *et al.* 1999; Houghton 2003). Reducing the uncertainty in these estimates is a trickle down process beginning with understanding the sources and implications of uncertainty at the smallest scale and incorporating it into ever-larger scales. Uncertainty analyses of process-based models such as SiB2.5 are part of this progression.

Uncertainty analyses are done in many ways, and the variety can be confusing. It is not uncommon to find the 'one at a time' (OAT) style approach used to equate step changes in parameters with uncertainty in results (Alapaty *et al.* 1997; Hallgren and Pitman 2000; Molders 2001). However, OAT style approaches and others like them have been shown to be incapable of capturing the net uncertainty of complex systems (Niyogi *et al.* 1999). Because of the complex nature of land-surface atmosphere modeling, where non-linear interactions are more the rule than the exception, it makes sense to view uncertainty in the context of the whole system rather than individual parts. Bayesian statistics provide a straightforward way to do this. As presented here, Bayesian statistics can be used to estimate the probability of a response given unknowns in the whole of the modeling system without assuming a true model and therefore a true response. Bayesian statistics are just starting to make headway in ecology (Ellison 1996; Ver Hoef 1996; Calder *et al.* 2003; Clark 2003; Wikle 2003), despite some early resistance (Dennis 1996), and are already frequently applied in atmospheric science (Kheshgi *et al.* 1999; Gurney *et al.* 2002) and hydrology (Beven and Binley 1992; Beven 1993; Freer and Beven 1996).

In this chapter, the predictive uncertainty of the SiB2.5 model is examined. Within the framework of Bayesian inference, predictive uncertainty can be understood as the probability that the actual response of a systemwill be predicted conditional on a set of circumstances. Building on the results from the last chapter, uncertainty bounds are calculated for the predicted fluxes of latent and sensible heat as well as for net ecosystem exchange. The predictive uncertainty of SiB2.5 is analyzed through time, on a monthly basis and for an entire year of simulations.

Several questions concerning the predictive uncertainty of SiB2.5 and of land surface models in general are addressed. They are:

1. Does predictive uncertainty vary with time? Does it vary diurnally? Does it vary seasonally? What does this tell us about the model? What does it tell us about the ecosystem we are studying?
2. Does adding net ecosystem exchange as an additional constraint reduce the predictive uncertainty of latent and sensible heat flux?
3. What are the potential impacts of predictive uncertainty for a land surface model such as SiB2.5?

4.2 Background

Uncertainty exists at every level of the modeling process, from the structure of a model itself, to the data that drives the model, to the observations the model is tested against (Kremer 1983). The sources of uncertainty are both known and unknown. Measurements often have known physical limits to precision; phenomena that occur below this level of precision while potentially important are not observable. For example, measurements of CO_2 concentration are only as precise as their calibration allows. Some measurements are also used under theoretical assumptions, such as tower flux data, where in many models it is assumed that observations of net radiation and fluxes calculated by eddy covariance are measuring precisely the same environment. When this assumption is violated, the result is uncertainty in the relationship between the two. Stochastic events can also cause variability in observations that may not be represented in models.

Initial conditions and environmental driver data are special forms of observations, ones that are used to run the model rather than to test it against. They suffer from many of the same sources of uncertainty. In addition, initial conditions and driver data are often subject to modifications, such as gap filling and estimation. Land surface models typically need continuous environmental driver data, such as air temperature, air pressure and wind speed. However, the measurements are often not continuous as a result of factors such as instrument failure and thus the gaps need to be filled. The gaps are filled with the best available estimate for the missing data, but nonetheless introduce uncertainty into data used by the model. Similarly, some initial conditions required to run land surface models may not be directly observed, such as soil moisture profiles, leaf area index or surface roughness length, and therefore must be estimated.

The structure of the model itself imparts an unknown level of uncertainty. All models operate under certain assumptions and the accuracy of these assumptions affects the level of confidence one should have in the resulting predictions. In land surface models such assumptions might be with regards to how a complex canopy is represented (e.g. big leaf or multi layer), which processes dominate others and on what time scales different processes operate. Further, as was shown in the previous chapter, there is typically not a unique parameter set that produces a paramount set of simulation results. Complex models are predisposed to exhibiting equifinality and parametric uncertainty tends to lead to greater predictive uncertainty (Franks 1998). Still another source of uncertainty is in the coding of the model itself; few equations can be solved analytically and choices are made about how processes are mathematically represented and how they are calculated through time. Taken as a whole, these various forms of uncertainty coalesce into the overall predictive uncertainty of the modeling system.

A methodology for calculating the predictive uncertainty of models has been developed by Beven and Binley (Beven and Binley 1992). The GLUE (Generalized Likelihood Uncertainty Estimation) methodology utilizes probabilistic likelihood measures to weight model predictions and calculate uncertainty bounds. The likelihoods represent the probability that a model simulation using a particular parameter set will be a good predictor of the system under investigation. GLUE is a methodology, and as such is a general system of conventions rather than a highly specific procedure. The requirements for GLUE are (Beven and Binley 1992): 1) a formal definition of a likelihood measure 2) an appropriate initial distribution of parameter values 3) a procedure for using likelihood weights in uncertainty estimation 4) a procedure for updating likelihood weights 5) a procedure for evaluating uncertainty and the value of additional data. Bayes theorem, which calculates conditional (or inverse) probability (James and James 1976) is used in GLUE to both calculate weights from the likelihoods and to update the likelihood weights with new information if it becomes available.

In the previous chapter, a number of the steps necessary for a GLUE style analysis were completed. A likelihood measure was defined (eqs. 3.5 and 3.6) and appropriate initial distributions of parameter values were determined (Table 3.1) fulfilling requirements 1 and 2. The likelihood measures and parameter distributions were then used to condition the model parameter space by running two sets of Monte-Carlo style analyses to identify the best performing parameter sets from 20,000 randomly generated sets. In the extension of that work to the uncertainty analysis presented here, the top ten percent of the simulations, as determined by likelihood value, are used to calculate the likelihood weights in requirement 3 of the GLUE procedure. How this was done is discussed in section 4.3 of this chapter.

While quite a number of studies utilizing the GLUE methodology have been carried out on hydrological models (Beven and Binley 1992; Romanowicz *et al.* 1994; Freer and Beven 1996; Franks *et al.* 1998; Brazier *et al.* 2000; Beven and Freer 2001; Brazier *et al.* 2001), relatively fewer have evaluated land surface-atmosphere models (Franks *et al.* 1997; Franks and Beven 1999; Franks *et al.* 1999; Schulz *et al.* 2001). This is not surprising, as the genesis of GLUE was in the discipline of hydrology. However, the methodology is well-suited to land surface-atmosphere modeling, since it implicitly handles non-linear interactions.

Franks and Beven (Franks and Beven 1997) examined the predictive uncertainty of the soil-vegetation-atmosphere-transfer (SVAT) scheme TOPUP with the GLUE methodology. TOPUP, a modified version of the hydrological model TOPMODEL, is a relatively simple SVAT designed primarily to predict evapotranspiration. In their analysis, they examined the predictive uncertainty of TOPUP with respect to daytime fluxes of latent heat in two different climates, the Amazon and Kansas, for two periods of time (one wet, one dry) each (in the Amazon, 17 and 73 day periods; in Kansas, 6 and 11 day periods). When conditioning the model on the individual time periods, they found that the predictive uncertainty bounds on latent heat flux were larger in an absolute sense for wetter periods than for drier periods, though in both cases the bounds were approximately 1/3 of the actual flux. When they conditioned on one period (wetter) and used the accepted parameter sets to calculate the predictive uncertainty of the second (drier), they found that the uncertainty in the predictions greatly increased. Their results suggest that when using a model to simulate different time periods that represent markedly different environmental conditions it is important to be aware that different process constraints might be active at different times and/or on processes of different durations. They also suggest that using as much data in the conditioning process as possible (more than they had available) might eliminate this problem and reduce predictive uncertainty.

In a follow on to the study mentioned above, Franks et al. (Franks *et al.* 1999) investigated whether including further constraints on the TOPUP SVAT model would reduce predictive uncertainty. In this analysis, they evaluated the same 73-day period from the Amazon, but also considered sensible heat flux (H) and ground heat flux (G).

They found that adding the additional constraints of H and G when conditioning the model parameter space helped to greatly reduce uncertainty. They suggested this was because energy balance is a bounded problem in the formulation of TOPUP, and the extra constraints of H and G helped to place appropriate limits on all three fluxes.

Schulz et al. (Schulz *et al.* 2001) investigated the predictive uncertainty of land surface fluxes with respect to increasing ambient [CO_2] using the GLUE methodology. In their study they used a simple photosynthesis model to evaluate what the impacts of predictive uncertainty are on predictions of latent heat flux and CO_2 flux in the presence of elevated CO_2 concentrations. They conditioned their model on the daytime fluxes of latent heat and CO_2 of a boreal agricultural field site in Sweden over the course of a two-week period in summer. They used two different scenarios to condition the model, one where six parameters were varied but the biochemical parameters of photosynthesis were fixed and one with 21 variable parameters that included the biochemical parameters of photosynthesis. They found that including more parameters in the conditioning of the model produced greater parametric uncertainty. This greater parametric uncertainty translated into greater predictive uncertainty when predictions of latent heat and CO_2 flux were made with elevated concentrations of CO_2. Their results suggest that it is important to be informed of the parametric uncertainty of a model that will ultimately be used to make predictions of the future.

The analysis of SiB2.5 presented in this chapter benefits from the results of these previous studies of land surface-atmosphere models while taking the analysis in a different direction. Unlike in other land surface-atmosphere studies using the GLUE methodology, the predictive uncertainty of a demonstrably more complex land surface-atmosphere model is evaluated for an entire year thereby capturing daily, monthly *and* seasonal patterns. This is important because the effects of different model parameterizations might not be evident on short time scales. For example, imagine a set of soil parameters that facilitated the drying out of the soil. Over short time scales they might have a negligible impact, however over longer time scales the soils would eventually dry out, causing a cascading series of effects. The seasonal patterns of uncertainty are related to model structure and logic, model sensitivity and to the relationship between deterministic models and comparatively stochastic measurements. How the different constraints of energy balance and the net ecosystem exchange of carbon affect predictive uncertainty in time are also examined. Because it was possible to evaluate the model for an entire year, it was also feasible to assess the appropriateness of conditioning model parameter space on all of the available data.

4.3 Methods

The uncertainty bounds, or prediction quantiles, were calculated based on the results from the sensitivity analysis of Chapter 3. In Chapter 3, a Monte Carlo style analysis was performed wherein 20,000 randomly generated parameter sets were ranked based on their capacity to simulate latent and sensible heat flux and net ecosystem exchange for each month of the year 1997 and for the entire year. Two criteria were used to evaluate the success of the simulations; the first criterion evaluated the ability of SiB2.5 to simulate latent and sensible heat flux (Eq. 3.5) the second criterion included net ecosystem exchange as an additional constraint (Eq. 3.6). The best performing parameter sets (top 10%) for each criterion were retained for each month and for the year as a whole, 28,000 in total. The sensitivity analysis in Chapter 3 was carried out using these top-performing sets, which were retained and used in this analysis.

Each retained parameter set was used to run the model in precisely the same manner as described in Chapter 3; for the year 1997, with a 6–year spin up and identical weather forcing. For each parameter set, hourly fluxes of sensible and latent heat as well as net ecosystem exchange were saved for the particular month or year associated with that parameter set. These hourly flux estimates were then used to calculate uncertainty bounds on the simulated output of the model.

The output distributions of the flux estimates for any given time step are not expected to consistently produce a normal distribution, therefore standard methods for calculating confidence intervals in most cases will not apply (Beven and Binley, 1992). Instead, the performance criteria (the likelihoods; L) from the sensitivity analysis were used. The general form of the likelihood criteria used in Chapter 3 is (Franks 1998):

$$L\left(\theta \mid \underline{y}, \sigma_\alpha^2, \sigma_\beta^2, \ldots\right) = \left(\left(\frac{\sigma_\alpha^2}{\hat{\sigma}_\alpha^2}\right)\left(\frac{\sigma_\beta^2}{\hat{\sigma}_\beta^2}\right)\ldots\right)^{-N} \tag{4.1}$$

Where:

$\theta =$ parameter set
$\underline{y} =$ observations
$\sigma_\alpha^2 =$ variance of errors for a particular variable
$\hat{\sigma}_\alpha^2 =$ minimum estimate of the variance of errors for the time period
$N =$ scaling, shaping factor to impose stringency/leniency

In the sensitivity analysis the scaling factor N was not implemented since the number of retained parameter sets was based on a percentage of the total number of parameter sets rather than on an imposed performance limit (See discussion in Chapter 3). Including N does not change the rank order of the results but does weight better simulations proportionately more. In this chapter, the likelihood values for each parameter set were recalculated using the shaping factor N = 1 to weight better simulations proportionately more. Equation 4.2, from Beven and Freer (Beven and Freer 2001), illustrates how Bayes equation is used with the calculated likelihoods to compute the posterior likelihood of each of the top 2000 simulations.

$$L\left[M\left(\theta_i\right)\right] = \frac{L_o\left[M\left(\theta_i\right)\right] L_T\left[M\left(\theta_i \middle| Y_T, Z_T\right)\right]}{C} \tag{4.2}$$

Where $L_o\left[M\left(\theta_i\right)\right]$ is the prior likelihood for the model $M\left(\theta_i\right)$ with parameter set θ_i; $L_T\left[M\left(\theta_i \middle| Y_T, Z_T\right)\right]$ is the likelihood measure calculated for the model with parameter set θ_i over time period T conditioned on input vector Y_T and observed variable vector Z_T; and C is a scaling constant such that the cumulative sum of the posterior likelihood, $L\left[M\left(\theta_i\right)\right]$, equals one. Because we used non-informative priors (uniform distributions for each parameter), $L_o\left[M\left(\theta_i\right)\right]$ is constant. The likelihood measures, $L_T\left[M\left(\theta_i \middle| Y_T, Z_T\right)\right]$, used in this chapter are the same as those used to calculate parameter sensitivity in Chapter 3 (Eq. 3.5 and 3.6)

Equation 4.3 (also from Beven and Freer 2001) illustrates how the prediction quantiles ($P\left(\hat{Z}_t < z\right)$; uncertainty bounds) are then calculated from the retained posterior likelihoods.

$$P\left(\hat{Z}_t < z\right) = \sum_{i=1}^{N} L\left[M\left(\theta_i\right) \middle| \hat{Z}_{t,i} < z\right] \tag{4.3}$$

Where $\hat{Z}_{t,i}$ is the value of variable Z at time t simulated by model $M(\theta_i)$ with parameter set θ_i. The computation of 90% uncertainty bounds for *each hourly time step* proceeds in this way: 1) predictions of latent and sensible heat flux and net ecosystem exchange from all two thousand runs are ranked from low to high flux, $\hat{Z}_{t,i}$ 2) likelihood values associated with each parameter set are maintained along with the fluxes as they are ranked, $L\left[M(\theta_i)\middle|\hat{Z}_{t,i} < z\right]$ 3) the likelihood values are then summed until 0.05 and 0.95 are reached $\sum_{i=1}^{N} L\left[M(\theta_i)\middle|\hat{Z}_{t,i} < z\right]$ 4) the flux at these points (*z*; 0.05 and 0.95) are the 5th and 95th prediction quantiles/uncertainty bounds for that time step. Uncertainty bounds such as these are probabilistic and represent the predictive uncertainty of the SiB2.5 model as conditioned on the input data, the observations, the parameter sets and the chosen likelihood measure (Beven and Freer 2001).

4.4 Results

The predictive uncertainty bounds were constrained by two multi-objective criteria, the model predictions fit to observed latent and sensible heat fluxes and the predictions fit to observed latent and sensible heat fluxes plus net ecosystem exchange. These multi-objective criteria were applied both on a monthly basis as well as for the entire year. The results are *hourly* sets of predictive uncertainty bounds for latent and sensible heat flux as well as net ecosystem exchange for the entire year of 1997 constrained both monthly and annually on the two multi-objective criteria.

The expectation, given a true model of the system, is that the predictive uncertainty bounds will be wide enough to encompass the observations and that the observations will fall in the center of the bounds (Beven and Freer, 2001). If the observations are not in the center of the bounds or are outside of the bounds entirely it can indicate several things. If observations are systematically above or below the uncertainty bounds during different parts of the day or during different parts of the year, that can be an indication of problems with model structure or logic. The parameters were given a large amount of freedom to vary and so systematic errors indicate an inability of the model to capture some process or environmental condition. In some instances, however, it can also indicate problems with the observations. For example, it has been observed that nighttime estimates of net ecosystem exchange are systematically under-observed because the lack of turbulence at night violates one of the main assumptions of the eddy covariance method (Goulden *et al.* 1996a; Lavigne *et al.* 1997; Law *et al.* 1999). When observations fall more randomly outside the predictive uncertainty bounds, not biased by time of day or day of year, it is an indication of variability in the measurements that the model cannot simulate.

Figure 4.1 shows example diurnal cycles of net ecosystem exchange, latent and sensible heat flux for two days in 1997, July 29 and October 23. This figure is merely an example to show the variability in observations and in predictive uncertainty bounds at two different times of the year. The predictive uncertainty bounds shown were constrained monthly with the NEE,LE,H criterion. Notice that the observations of net ecosystem exchange fall within the uncertainty bounds in both time periods and that they are quite well centered (Figure 4.1a,d). For sensible heat flux (Figure 4.1c,f), the observations are also quite well centered within the uncertainty bounds in July, however in October the observations fall well above the uncertainty bounds during the daytime. For latent heat flux there is more variability during both time periods (Figure 4.1b,e). In both July and October, the model appears to be unable to simulate the relatively higher afternoon latent heat flux. Notice also how the magnitude of the uncertainty bounds varies with time of day (or magnitude of flux). The predictive uncertainty bounds are very narrow during the nighttime for both latent and sensible heat flux, indicating SiB2.5 does not simulate much variability in heat fluxes at night.

4.4.1 *Hourly predictive uncertainty*

There are patterns in both the observations and the hourly predictive uncertainty bounds that occur within a month. In October, for example, leaf off occurs in the days around October 10th. After this time, both the observations and uncertainty bounds for net ecosystem exchange change in magnitude, adjusting to the loss of photosynthesizing vegetation (Figure 4.2a). During the time period just before and after leaf-off, observations of nighttime net ecosystem exchange fall well below the predictive uncertainty bounds, indicating systematic under prediction of nighttime NEE during this short time period. The daytime observations, on the other hand either fall inside the predictive uncertainty bounds or just below (indicating a slight tendency to over predict daytime NEE). SiB2.5 does not explicitly treat heterotrophic respiration from litter so perhaps this under prediction of nighttime NEE is related to a first flush of respiration after leaf fall. In other months, such as July, the relationship between observations and the predictive uncertainty bounds is more stable (Figure 4.2b). There are no obvious patterns of over or under predicting net ecosystem exchange and the observations fall agreeably within the predictive uncertainty bounds, indicating good model structure and logic during this time period

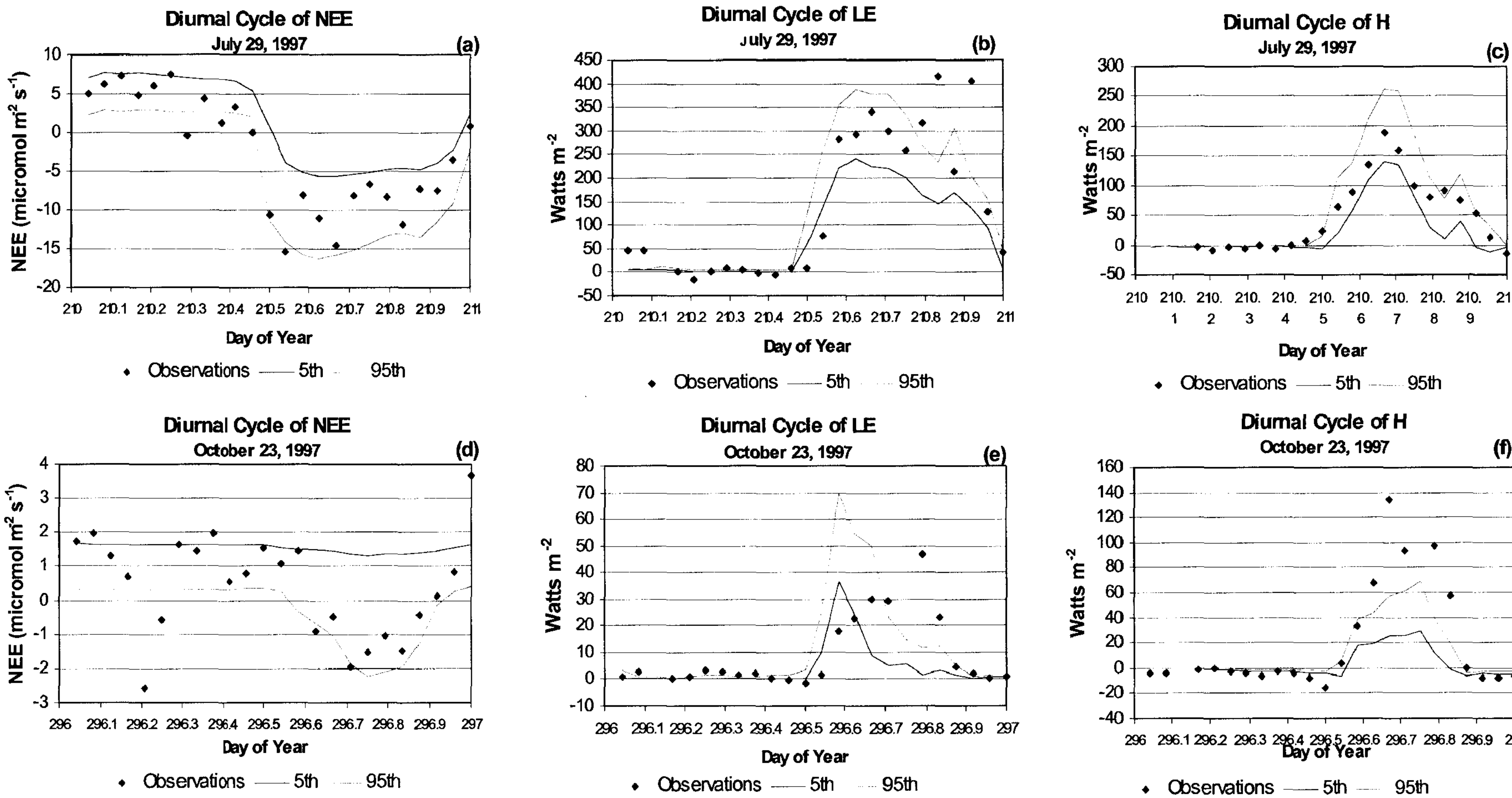


Figure 4.1 Diurnal cycles of observations and predictive uncertainty bounds for July 29 (a,b,c) and October 23 (d,e,f), 1997. Predictive uncertainty bounds were constrained with the NEE,LE,H criterion and parameter sets were conditioned monthly.

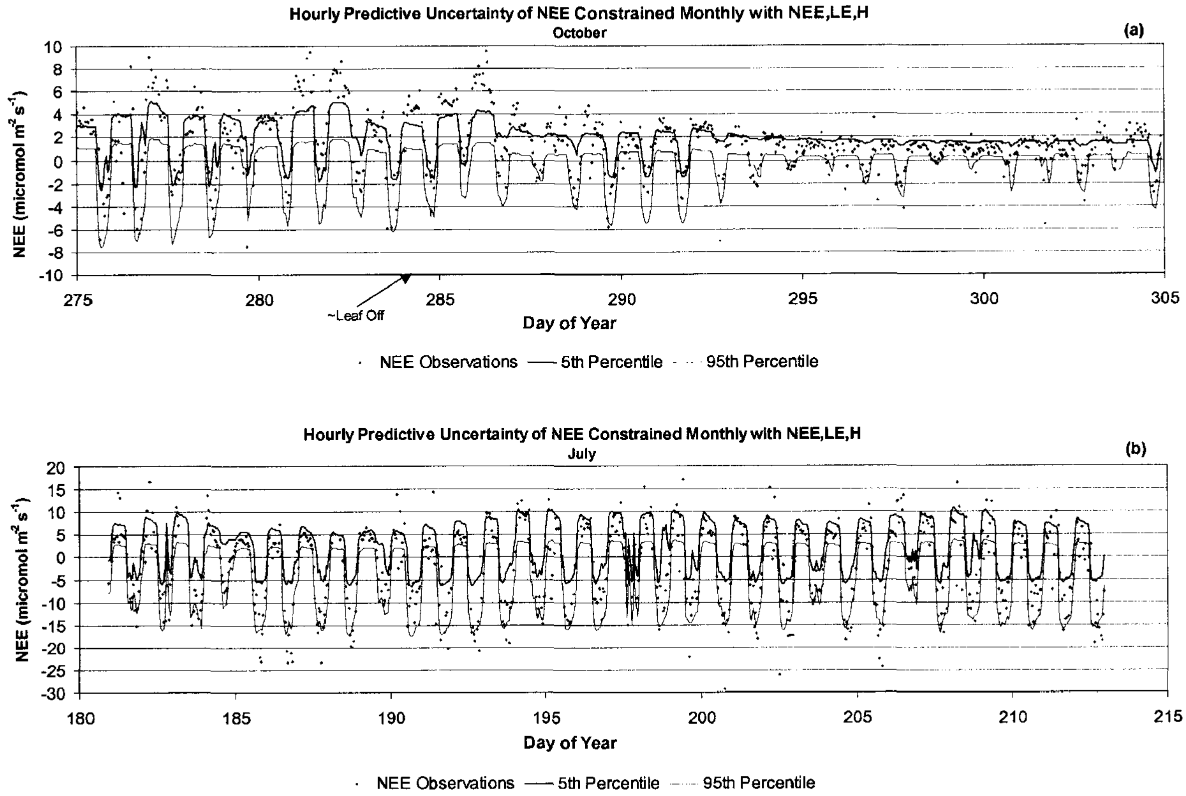


Figure 4.2 Hourly predictive uncertainty bounds for net ecosystem exchange constrained monthly on the NEE,LE,H criterion **a)** October, 1997 **b)** July, 1997.

4.4.2 *Monthly mean diurnal predictive uncertainty*

Monthly mean diurnal plots provide a useful summary of the results. Figures 4.3 and 4.4 present the monthly mean diurnal uncertainty in simulations of net ecosystem exchange and latent and sensible heat flux constrained monthly on LE, H and NEE,LE,H. These plots show broad patterns in the fit of the predictive uncertainty bounds to the observations. Considering the results already presented for July and October and the NEE,LE,H criterion, which relate directly to Figure 4.4, we can see that the monthly mean diurnal plots do a reasonable job in capturing the patterns observed at a finer temporal scale.

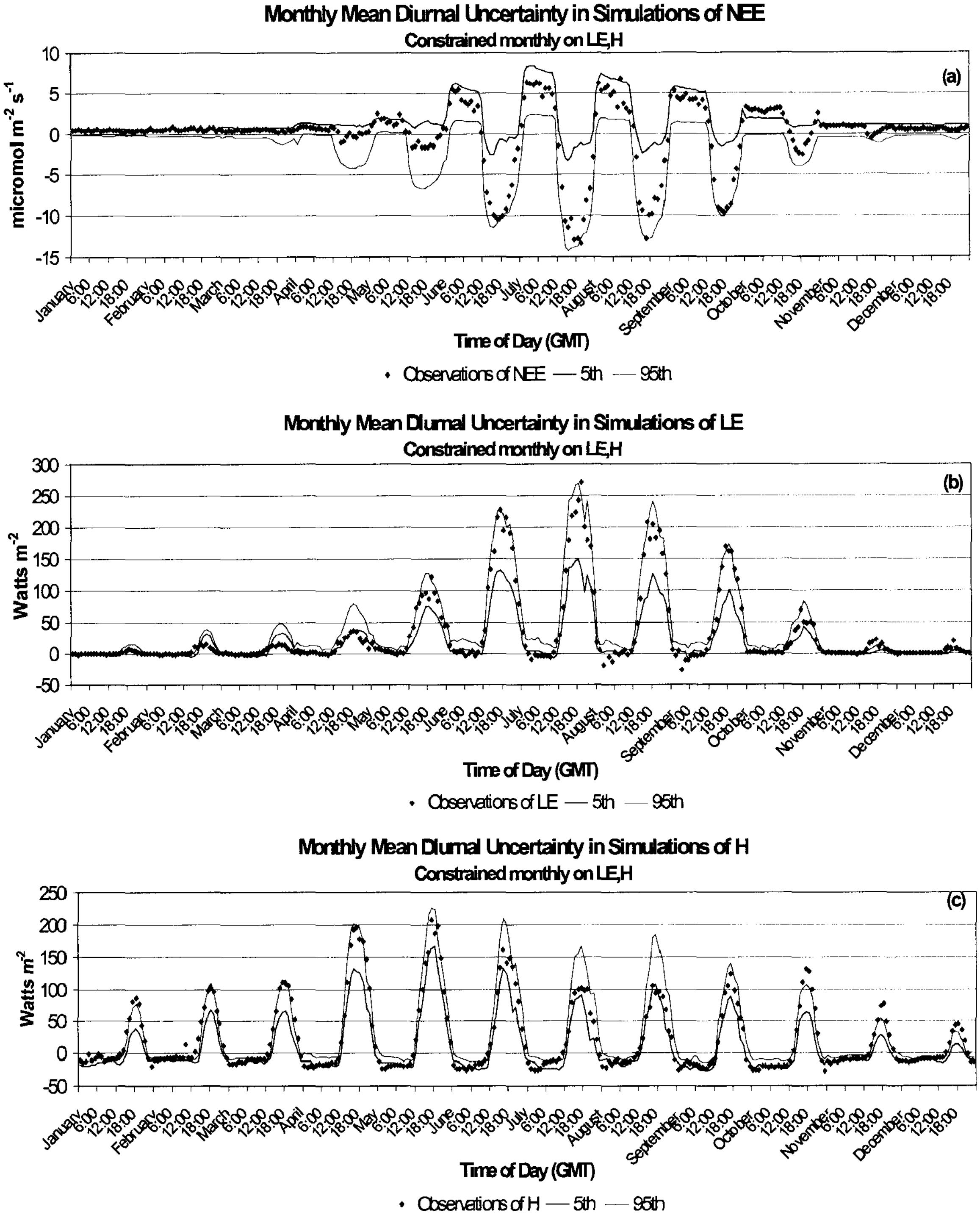


Figure 4.3 Monthly diurnal mean uncertainty in simulations of **a)** net ecosystem exchange, **b)** latent heat flux and **c)** sensible heat flux constrained monthly on LE,H. Gray diamonds represent the monthly diurnal mean of the observations, the blue solid line represents the monthly diurnal mean of the 5% predictive uncertainty bound, the red solid line represents the monthly diurnal mean of the 95% predictive uncertainty bound. The diurnal cycle for each month is shown.

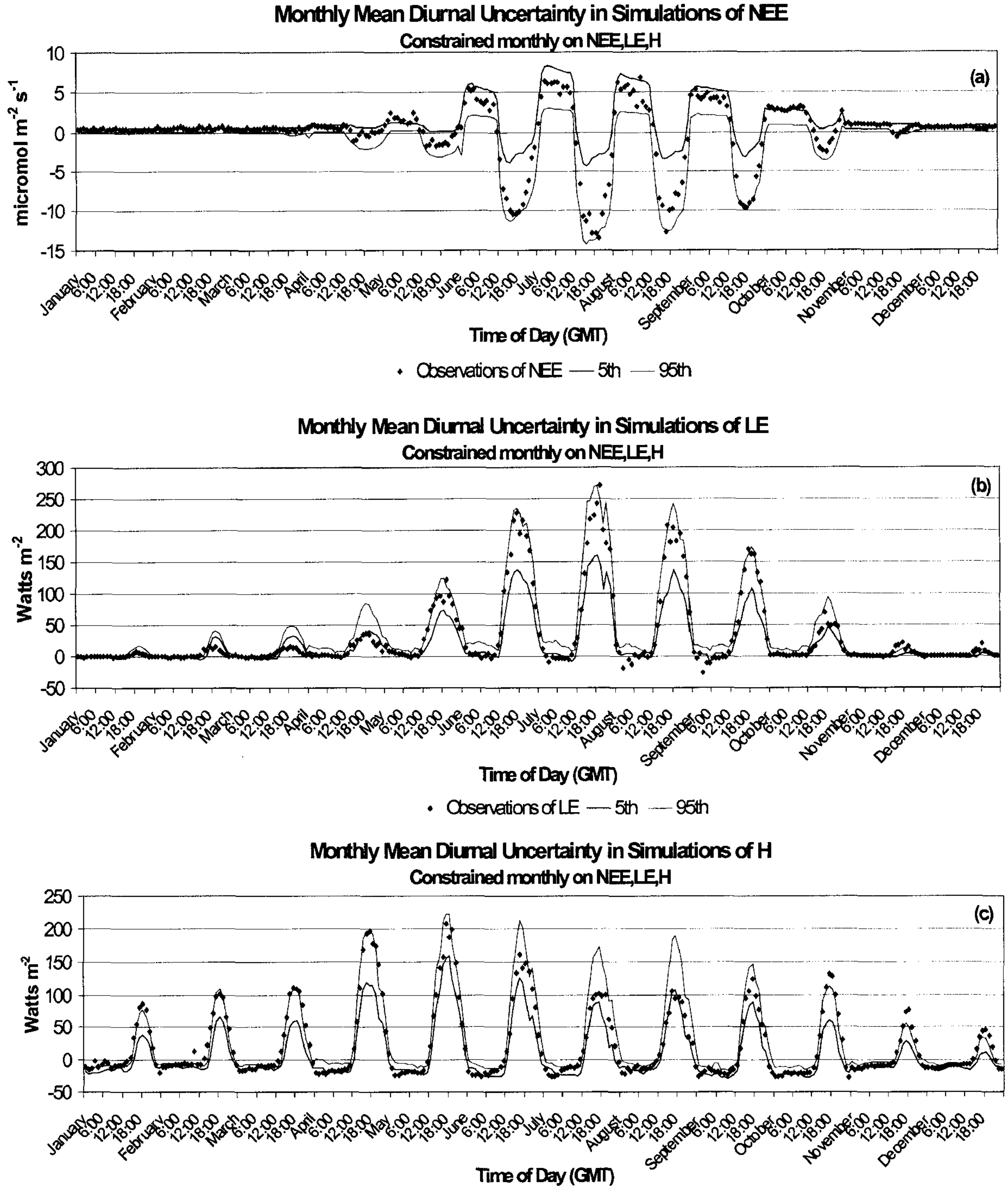


Figure 4.4 Monthly diurnal mean uncertainty in simulations of **a)** net ecosystem exchange, **b)** latent heat flux and **c)** sensible heat flux constrained monthly on NEE,LE,H. Gray diamonds represent the monthly diurnal mean of the observations, the blue solid line represents the monthly diurnal mean of the 5% predictive uncertainty bound, the red solid line represents the monthly diurnal mean of the 95% predictive uncertainty bound. The diurnal cycle for each month is shown.

The two constraints, LE,H and NEE,LE,H, resulted in very similar predictive uncertainty bounds for both the latent heat flux and sensible heat flux, with the exception of sensible heat flux in October, where more observations fell within the uncertainty bounds with the inclusion of net ecosystem exchange in the constraint. For both latent and sensible heat flux, rarely are the observations in the center of the predictive uncertainty bounds. For daytime latent heat flux, the observations appear most centered in May and August. For daytime sensible heat flux the observations appear most centered in June and September.

Daytime latent heat fluxes show a tendency to be systematically over-predicted in the winter and spring months (Dec,Jan,Feb,Mar,Apr) and under-predicted in the summer months of June and July. Evening latent heat fluxes appear to be systematically over-predicted in most months, though as we have seen the nighttime uncertainty bounds are very narrow. Daytime sensible heat fluxes are systematically under predicted in the winter, spring and fall months (Jan,Feb,Mar,Apr,Oct,Nov,Dec) and show a tendency toward over-prediction in the summer (Jul,Aug). For nighttime sensible heat fluxes it is less clear, though it does appear that early evening sensible heat fluxes tend to be over-predicted and that nighttime sensible heat fluxes in April, May and October are over-predicted. Evaluation of the percentage of hourly observations above, within and below the hourly predictive uncertainty bounds generally supports this interpretation (Tables 4.1 and 4.2), though it is difficult to separate out daytime versus nighttime fluxes in the percentages. It is surprising to see that for both latent and sensible heat flux, less than 50% of the observations fall within the uncertainty bounds. The absolute percentage varies by season with the most observations falling within the predictive uncertainty bounds in the summer months and the fewest in the winter.

Table 4.1 Percentage of observations above, within and below 90th percentile predictive uncertainty bounds for latent heat flux, LE,H constrain and NEE,LE,H constraint, conditioned monthly.

	Percentage of Observations Above, Within or Below Predictive Uncertainty Bounds for LE					
	LE,H			***NEE,LE,H***		
	% Above	***% Within***	***% Below***	***% Above***	***% Within***	***% Below***
January	43	19	38	42	19	39
February	53	11	36	52	12	36
March	44	13	43	45	13	42
April	21	19	60	20	19	61
May	40	33	27	43	30	27
June	29	35	36	28	36	36
July	28	36	36	28	36	36
August	32	33	35	33	32	35
September	26	29	45	26	28	46
October	32	29	39	30	31	39
November	44	20	36	44	21	35
December	54	15	31	55	16	29

Table 4.2 Percentage of observations above, within and below 90th percentile predictive uncertainty bounds for sensible heat flux, LE,H constrain and NEE,LE,H constraint, conditioned monthly.

	Percentage of Observations Above, Within or Below Predictive Uncertainty Bounds for H					
	LE,H			***NEE,LE,H***		
	% Above	***% Within***	***% Below***	***% Above***	***% Within***	***% Below***
January	48	15	37	48	16	36
February	48	18	34	48	18	34
March	48	18	34	49	18	33
April	29	31	40	29	31	40
May	29	28	43	30	28	42
June	23	39	38	23	39	38
July	20	44	36	21	44	35
August	25	33	42	25	34	41
September	19	39	42	20	40	40
October	38	22	40	38	27	35
November	43	20	37	44	20	36
December	50	18	32	51	21	28

Energy balance is enforced in SiB2.5 as it is in most single point models. In many months, where one heat flux shows systematic over-prediction, the other shows systematic under-prediction. This analysis does not include ground heat flux as an additional constraint, which might have helped moderate the swings in over and under-prediction of sensible and latent heat (Franks *et al.* 1999). However, ground heat flux is a point measurement at eddy covariance sites, so it is not clear that it would have been appropriate to add it as an additional constraint. There are also finer features presented in monthly mean diurnal plots specific to the fluxes. For example, the model appears to miss high afternoon fluxes of latent heat in July (Figure 4.4b, 4.5a). For sensible heat flux, the model misses an early evening dip in sensible heat flux (Figure 4.4c, 4.5b).

Observations of net ecosystem exchange appear to be centered within their predictive uncertainty bounds more often than observations of latent and sensible heat fluxes were (Figures 4.3a, 4.4a). August and October show the most centered observations for daytime fluxes for both constraints: for the NEE, LE, H constraint, daytime observations were also well centered in May. For nighttime fluxes, the observations are most centered for both constraints during the summer months of June, July and August. In the spring and fall (Apr, May, Oct, Nov) nighttime fluxes tend to be under-predicted. In the beginning and end of the leaf-on period (Jun, Sep), daytime fluxes show a tendency towards under-prediction.

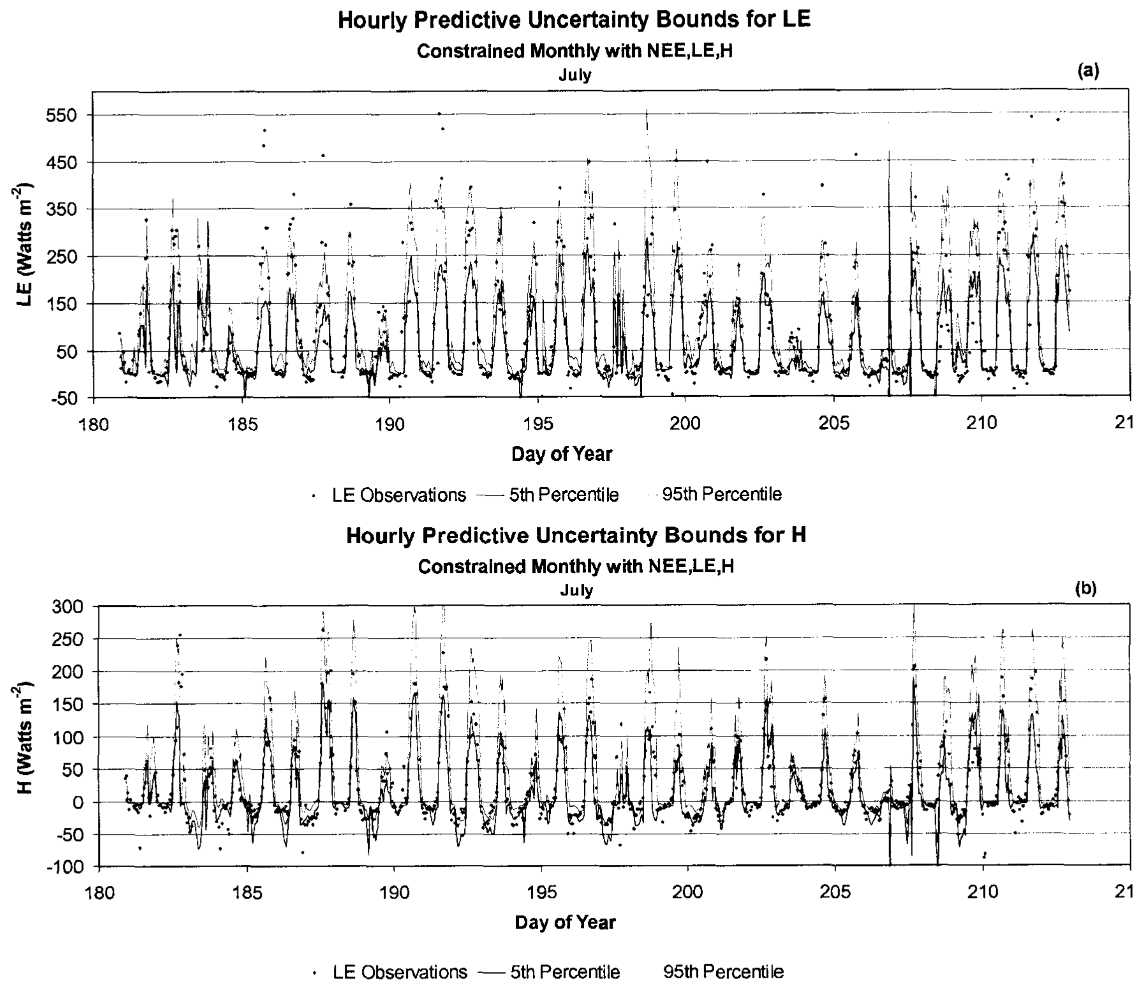


Figure 4.5 Hourly predictive uncertainty bounds for latent and sensible heat flux constrained monthly on LE,H for July 1997 a) latent heat flux b) sensible heat flux.

There are more obvious differences between the predictive uncertainty bounds for the two constraints (LE,H and NEE,LE,H) with respect to observations of net ecosystem exchange than there were for observations of latent and sensible heat flux. April, May and October show the most difference, with the predictive uncertainty bounds, both the 5^{th} and 95^{th} percentile, being narrower in April and May for the NEE,LE,H constraint and the 5^{th} percentile uncertainty bound for October being narrower in the nighttime for the LE,H constraint. The narrower uncertainty bounds in May result in more centered observations during the daytime and fewer observations actually falling within the uncertainty bounds (Table 4.3). The summer months (Jun,Jul,Aug), which were time periods shown in Chapter 3 to have a large number of common simulations between the two constraints, show the least difference between the two constraints. This difference will be discussed further in a later section.

Unlike the fluxes of latent and sensible heat, the percentages of NEE observations within the uncertainty bounds are quite high. For the LE,H constraint, all months except October have greater than 50% of the observations within the predictive uncertainty bounds. For the NEE,LE,H constraint January, February, May and December fall below 50% of the observations within the bounds, indicating that fewer observations fall within the predictive uncertainty bounds when the NEE,LE,H constraint is applied than when the LE,H constraint is applied. Overall these results show that the structure and logic in SiB2.5 for the simulation of diurnal to seasonal variations in net ecosystem exchange is quite good.

Table 4.3 Percentage of observations above, within and below 90^{th} percentile predictive uncertainty bounds for net ecosystem exchange, LE,H constrain and NEE,LE,H constraint, conditioned monthly.

	Percentage of Observations Above, Within or Below Predictive Uncertainty Bounds for NEE					
	LE,H			***NEE,LE,H***		
	% Above	***% Within***	***% Below***	***% Above***	***% Within***	***% Below***
January	8	56	36	22	39	39
February	6	57	37	18	46	36
March	4	76	20	21	51	28
April	7	74	19	13	58	29
May	6	74	20	10	48	42
June	20	63	17	21	59	20
July	18	71	11	19	68	13
August	18	61	21	20	57	23
September	15	66	19	18	60	22
October	5	40	55	9	51	40
November	3	74	23	17	56	27
December	3	87	10	29	39	32

4.4.3 *Reducing predictive uncertainty with additional constraints*

As expected, adding in the additional constraint of net ecosystem exchange when conditioning the model on different parameters sets reduces the predictive uncertainty of net ecosystem exchange (Figure 4.6a,b). However, the reduction in uncertainty is not constant through time. The largest reductions in predictive uncertainty occur during the spring in April and May. In the last chapter, it was shown that when the model was constrained with only the LE,H criterion in the month of April, very unusual parameter sets resulted. The hypothesis was that the treatment of melting snow in SiB2.5 was inadequate, causing unusually high spikes in heat fluxes and therefore inappropriate parameter sets were selected to compensate for this problem. The addition of net ecosystem exchange as a constraint resulted in more biologically realistic parameter sets. The reduction in the parametric uncertainty resulted in the reduction of predictive uncertainty shown here. Similarly, premature greening up of the canopy in May due to temporal interpolation between monthly mean NDVI composites resulted in unrealistic parameter sets when only the LE,H criterion was considered. The inclusion of net ecosystem exchange again reduced parametric uncertainty leading to reduced predictive uncertainty. Predictive uncertainty in net ecosystem exchange was also reduced in the winter months with the addition of NEE to the constraint. However, during the rest of the year the reduction is less dramatic.

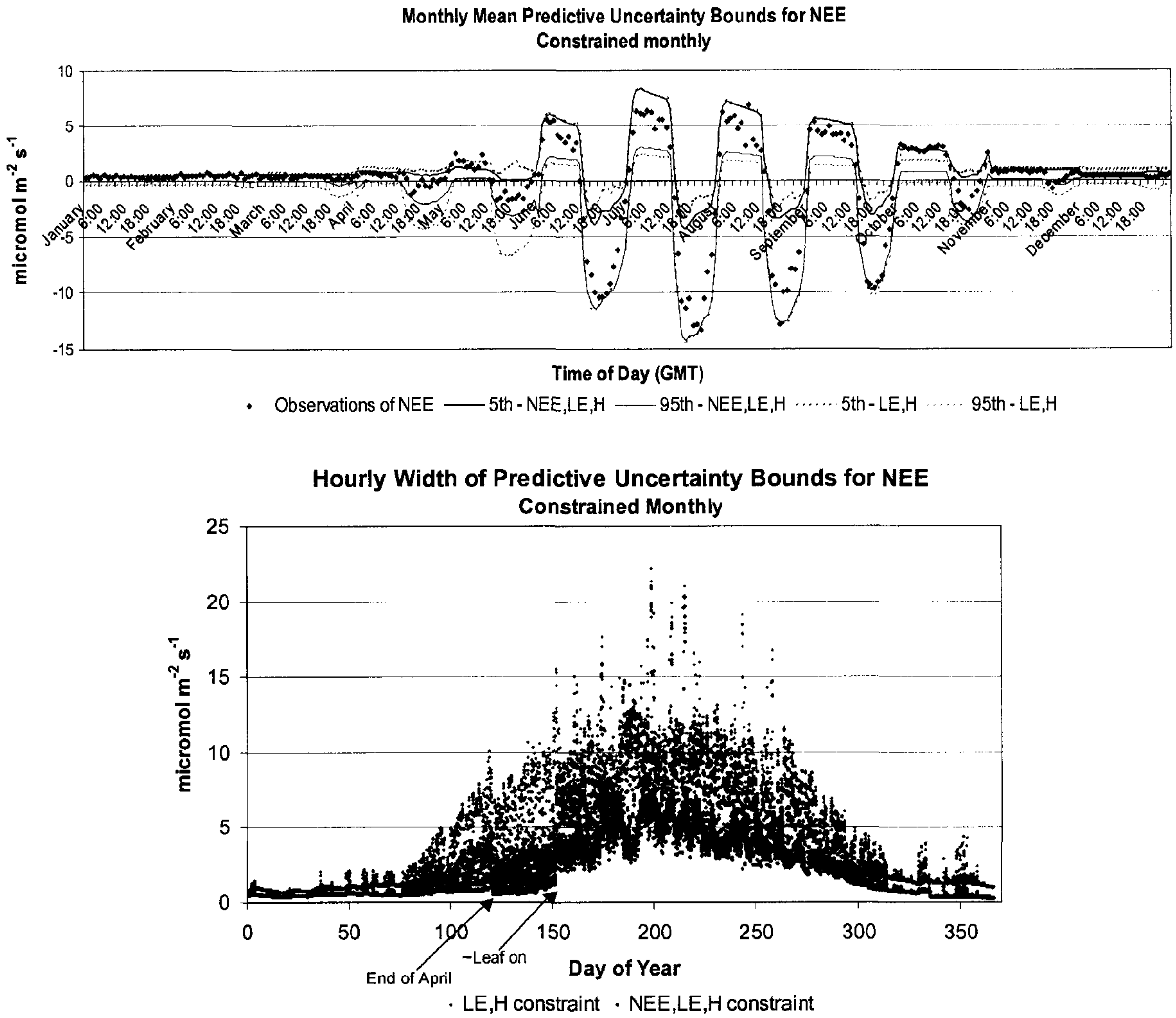


Figure 4.6 a) Monthly mean predictive uncertainty of NEE constrained monthly on LE,H (dashes) and NEE,LE,H (solid line) **b)** Hourly widths of the predictive uncertainty bounds constrained monthly for NEE under two constraints, LE,H (red) and NEE,LE,H (blue).

For the sensible and latent heat flux there is no appreciable improvement in predictive uncertainty with the additional constraint of net ecosystem exchange (not shown). This is somewhat surprising since one might have expected a reduction at least in latent heat flux given the coupling of transpiration and carbon assimilation in SiB2.5.

4.4.4 *Annual versus monthly constraints*

In addition to the two constraints of LE,H and NEE,LE,H that were used to condition the parameter space for the uncertainty analysis, two time constraints were evaluated. Predictive uncertainty was calculated for each flux for each data constraint by constraining the parameter space either monthly (presented thus far) or annually. For the monthly constraint model parameters were allowed to vary in time, for the annual constraint they were not. In some cases, such as for the month of April, the monthly constraint selected for biologically unrealistic parameter sets due to model structural problems (see discussion in Chapter 3). In other cases, the monthly constraint allowed parameters such as VMAX to vary, which is more biologically realistic. Figure 4.7 shows the monthly mean predictive uncertainty for each flux, constrained on the NEE,LE,H criterion monthly and annually.

There are differences between constraining monthly or annually but they don't fit a single pattern for the different fluxes. For net ecosystem exchange, overall the daytime fluxes were better captured by the monthly constraint and the nighttime fluxes by the annual constraint. This is reflective of the timescales on which these two processes operate in nature and in SiB2.5. Ecosystem respiration in SiB2.5 is calculated based on the annual assimilation of carbon by the vegetation, which is a reasonable assumption at WLEF for the year 1997 (discussed in detail in Chapter 2). The assimilation of carbon by vegetation in SiB2.5 is not restricted annually and proceeds as the environment and condition of the canopy dictate.

For latent and sensible heat flux, in general the monthly parameter adjustment produced predictive uncertainty bounds that were better able to capture the observations than the annual constraint. As with daytime net ecosystem exchange, fluxes of latent and sensible heat are more actively coupled to the environment at shorter temporal periods. In the case of latent and sensible heat flux, the flux is determined by the status of the landscape (wet or dry) as well as the atmospheric conditions (vapor pressure deficit, cloudiness, wind speed and direction, etc..). The time periods on which these fluxes are dependent are thus shorter than those for nighttime net ecosystem exchange and the monthly predictive uncertainty bounds are therefore more appropriate. For latent heat flux, the monthly predictive uncertainty bounds are narrower than the annual predictive uncertainty bounds in all months except July and August, similar to net ecosystem exchange. For sensible heat flux, the monthly predictive uncertainty bounds are equivalent or just narrower than the annual bounds in all months except April, May and June where the monthly uncertainty bounds are not necessarily much narrower than the annual bounds, but have shifted upward in a better fit to the observations.

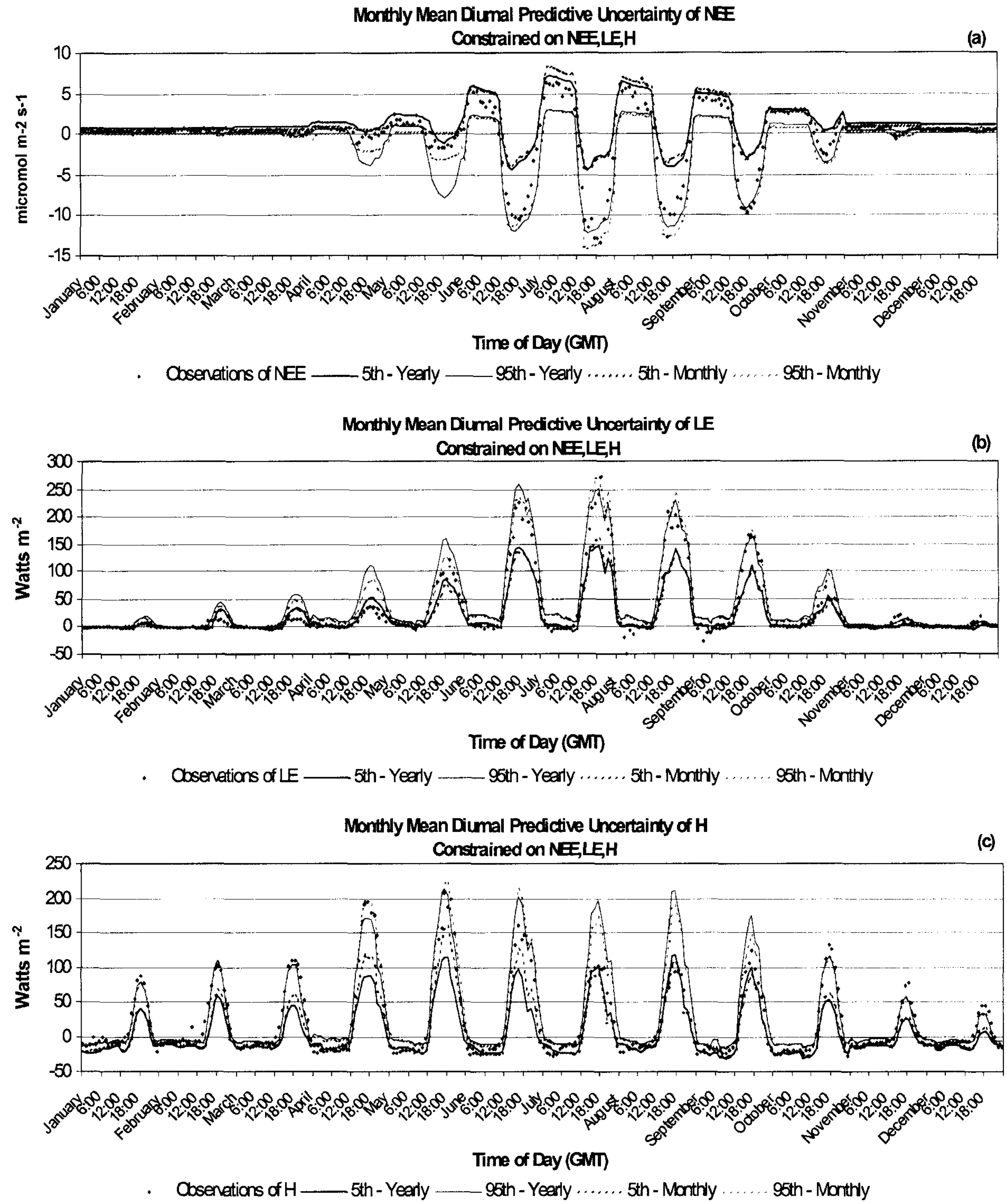


Figure 4.7. Monthly mean diurnal predictive uncertainty constrained yearly (solid line) and monthly (dashed line) on NEE,LE,H for **a)** NEE, **b)** LE and **c)** H.

4.5 Discussion

An analysis such as the one presented here is sensitive to the choices that define it. If this analysis were conducted with a different time period of data, in another geographical location or if other data from the same site were used to constrain the model the conclusions might have been be quite different. Further, the choice of which parameter sets to retain and even the choice of goodness of fit criteria will affect the results. For example, in this analysis the scaling factor N was set to 1. There is a direct correlation between the width of the predictive uncertainty bounds and the choice of N, the larger N is the narrower the uncertainty bounds will be (Freer and Beven 1996; Franks and Beven 1997). However, given that a truly generalized analysis is never possible and that these choices were explicitly made, the results that were obtained here can at least inform the expected uncertainty in simulated CO_2 arising from parameter error in SiB. This type of information is very useful for inverse modeling.

All land surface models must necessarily be simplifications of complex land surface–atmosphere interactions. If a model represented all of these complexities faithfully and highly accurate observations were made, we would expect, regardless of the constraints imposed, that the observations of the simulated quantity would fall within the center of the predictive uncertainty bounds. Because SiB2.5 was not developed to represent every aspect of the system and because the observations were not neatly centered within the uncertainty bounds, there is much to learn from an uncertainty analysis such as this. The observed biases, the responses to different constraints and the temporal variability discerned in this analysis provide insights on the ecosystem at WLEF, the structure and logic of the model and most importantly how these two intersect.

In the plots of individual days, months and monthly diurnal means, biases in the simulated fluxes as compared to the observed fluxes were observed. Because the parameters were given a large amount of freedom to vary, systematic errors indicate either an error in the structure or logic of the model that precludes the representation of a process or environmental condition or systematic errors in the observations themselves. The high percentage of observations that fell within the predictive uncertainty bounds for net ecosystem exchange indicates that the structure and logic of the model for this flux is quite good, at least at diurnal timescales. There were some observed biases however. Over and under estimation of daytime net ecosystem exchange in times of rapid landscape change (spring and fall) was associated with the lack of high temporal precision in the time varying parameterization of the model, especially leaf area index and f_{PAR}. Biases in the nighttime net ecosystem exchange, which were observed in some months, could be related to the simplified respiration code that responds to temperature, moisture and the annual total assimilation of carbon, but not the daytime uptake or recent input of litter.

Far fewer of the observations of sensible and latent heat flux fell within the predictive uncertainty bounds and biases were more obvious and persistent than for net ecosystem exchange. There are several possible reasons why heat fluxes are more problematic then net ecosystem exchange for SiB2.5. As mentioned previously, energy balance is enforced in most single point land surface-atmosphere models, however observations of the components of net radiation made in the field rarely close the surface energy budget. Further, despite careful rules to ensure that the flux measurements are representative of the heterogeneous landscape at WLEF, turbulence and landscape variability that is not characterized in SiB2.5 could be playing a significant role in the variability of the fluxes. For example, Ewers et al (Ewers *et al.* 2002) found large differences in the transpiration of four different forest types (northern hardwoods, conifer, aspen/fir, forested wetlands) found in the WLEF area; in the single point version of SiB2.5 a single functional type represents the WLEF area, mixed forest. Shifts in wind direction at certain times of day or days of year could produce fluxes dominated by one forest type or another that SiB2.5 is not intended to reproduce. If these sources of variabilitys are persistent, then the mean field parameterization of SiB2.5 must be considered incorrect for this site. It is possible that predicting net ecosystem exchange is less problematic because it is less variable across the different vegetation assemblages at WLEF than heat fluxes are and therefore mean-field parameterizations and functional groups are adequate representations.

Adding the additional constraint of net ecosystem exchange did not reduce the predictive uncertainty of latent and sensible heat flux. Given the coupling of transpiration and carbon assimilation in SiB2.5, this is somewhat surprising. It is possible that the greater variability of the heat fluxes mentioned above prevented the additional constraint from having much of an effect. We do know from the last chapter, however, that the additional constraint of NEE did reduce the parameter sets to more biologically meaningful groupings. As expected, the additional constraint did reduce the predictive uncertainty of net ecosystem exchange, although not across the board. The least amount of reduction in predictive uncertainty occurred in the summer months when net ecosystem exchange is the most coupled to the heat fluxes. During the winter months, when net ecosystem exchange is less coupled to the heat fluxes, there was a proportionally larger reduction in predictive uncertainty. The largest reduction, however, occurred during the spring and fall months, when rapid changes in the landscape (thaw/freeze/leaf-on/leaf-off) and shifts from a coupled to a less coupled state increased the power of an additional constraint.

It is common wisdom that the more data a model is conditioned and constrained with and the longer the period of time, the better and more robust the results will be. Depending on the temporal scale of variability of the quantity being predicted and the ability of the model to match that scale, this may or may not be true. The results presented here on constraining predictions annually versus monthly show that some quantities, such as nighttime net ecosystem exchange, were on average better constrained annually, whereas other quantities that show more variability, such as latent and sensible heat flux, were not. The results are partially dependent on the scale of variability represented in SiB2.5. For example, ecosystem respiration is an annually bounded flux in SiB2.5 that varies only with soil temperature and moisture. The assimilation of carbon and latent and sensible heat flux are more dependent on shorter-term environmental forcing in SiB2.5 and thus the monthly constraint was more appropriate. Constraining on shorter time periods also makes up for a lack of flexibility and variability in the model through adjustments in the parameterization. Understanding these compensations is important for the formulation of robust predictive models, because while some will be ecologically meaningful, others will not.

How is it possible to use this information on predictive uncertainty? Clearly one use is to investigate the sources of observed bias and correct the model and/or observations for deficiencies in structure or logic. Another use would be to incorporate the uncertainty into the modeling itself. In the analysis presented here there was clearly variability in the real world that was not captured by the model and may in fact not be easily calculated or parameterized. The addition of variability into the model predictions based on the level of observed uncertainty might be possible, particularly when ensemble runs such as these are impractical. Additionally, larger scale atmospheric models often must parameterize the uncertainty due to the variability in land surface fluxes in vary basic ways. An analysis such as this provides more robust estimates of predictive uncertainty as it varies through time and quantifies the expected model-data mismatch for inversion studies.

The analyses presented thus far in Chapters 2 and 3 were based on a single point model driven by observed weather. This kind of off-line testing is good for analyzing the sensitivity and uncertainty of model predictions with respect to model structure, logic and parameterization. In the next chapter a version of the SiB model coupled the regional atmospheric model RAMS (Regional Atmospheric Modeling System) is used to investigate the effect of the representation of spatial heterogeneity on regional carbon fluxes, providing a more integrated picture of the effects of land surface characterization on modeled carbon fluxes.

CHAPTER 5

SENSITIVITY OF REGIONAL CARBON FLUX TO LAND SURFACE HETEROGENEITY

5.1 Introduction

A great deal of progress has been made in recent years characterizing the carbon cycle at a variety of scales. In particular both "bottom up" (the use of fine scale process level understanding and spatial information to predict patterns at larger scales) and "top down" (the analysis of large-scale atmospheric patterns to infer process) approaches to carbon cycle science have improved substantially. However, at the crucial intermediate scale where these two approaches intersect and where direct comparison of the two approaches is possible, the progress has been a bit slower. There are many reasons for this. Intermediate scales are roughly at the limit of what can be explicitly modeled and what has to be parameterized. Restrictions on representation are often made based on computational limitations and available data as much as on what would be scientifically most appropriate. Further, comprehensive regional field measurement campaigns have been prohibitively expensive in the past and, while a handful of such experiments have taken place across the globe in recent years (FIFE (Sellers *et al.* 1992b), BOREAS (Sellers *et al.* 1995a), HAPEX-Sahel (Prince *et al.* 1995), LBA (Avissar *et al.* 2002)) they characterize only select environments. Thus both new kinds of measurements and new methodologies to interpret those measurements are sorely needed.

Over the last decade, a global network of towers measuring fluxes of CO_2, latent and sensible heat has been established (FLUXNET; (Baldocchi *et al.* 2001)). At scales of meters to several square kilometers, measurements from these sites along with mechanistic modeling have been used to quantify ecosystem fluxes of CO_2, latent and sensible heat and to explore the physical processes that govern them (see previous chapters). To extrapolate these measurements in space and make them more commensurate with top down approaches, techniques such as relating fluxes to spectral indices (Veroustraete *et al.* 1996; Rahman *et al.* 2001), extrapolating models fit to flux data across the landscape with spatial datasets (Oechel *et al.* 2000; Vourlitis *et al.* 2000), scaling tower flux measurements across the land surface by biome (Valentini *et al.* 2000) or relating tower flux measurements to aircraft measurements and scaling by ecosystem type (Mahrt and Vickers 2002) are used. However, there is not as yet a reliable way to evaluate these estimates in both space and time.

Top down approaches, such as global inversions, which use seasonal and interannual patterns of $[CO_2]$ from the global flask network to infer sources and sinks of CO_2 in the global biosphere, have been in use for many years (Tans *et al.* 1990; Ciais *et al.* 1995; Fan *et al.* 1998; Gurney *et al.* 2002). They have provided the scientific community with challenging problems to consider, for example the issue of the 'missing sink' of carbon (Ciais *et al.* 1995; Fan *et al.* 1998). However, it has been difficult to evaluate the various results since it has also been problematical to confirm large-scale source/sink patterns of CO_2 on the land surface.

Recently transport models have been applied at finer scales in an attempt to further resolve the patterns observed at larger scales. Kjellstrom et al. (Kjellstrom *et al.*

2002) investigated summertime CO_2 transport in Europe and Siberia. They used an off-line regional transport model with land surface fluxes prescribed from a land surface model at 1 x 1 degree resolution. They found that both synoptic meteorological variability and the heterogeneity of CO_2 fluxes make single concentration measurements difficult to interpret. They also found that it was difficult to determine an appropriate atmospheric level to capture a 'regional footprint' of CO_2 flux because of this variability. Geels et al. (Geels *et al.* 2004) investigated the high frequency variability of continental CO_2 concentration fields over Europe and North America using two different scales of land surface flux resolution, 1 x 1 and 3 x 3 degree resolutions, for an entire year at both monthly and daily temporal resolutions with a regional transport model. They found that high temporal resolution simulations were necessary to capture the high frequency variation in the continental CO_2 concentration record, but that the difference between land surface resolutions had less of an impact. Kaminski et al. (Kaminski *et al.* 2001) investigated the effects of surface flux resolution on synthesis inversions. Synthesis inversions use atmospheric concentrations of CO_2 to produce fluxes of CO_2 for different regions of the globe (basis regions). By increasing the number of basis regions, they were able to determine that aggregation of basis regions can produce flux estimate errors on the order of the fluxes themselves. However, they suggest it is not enough to simply increase the number of basis regions, but that a level of heterogeneity between regions must be maintained. It is also not always feasible to increase the number of basis regions, since it can increase computational requirements considerably. Though these three studies were performed at much higher resolutions than typically found in global transport and inversion studies, they are still quite coarse when compared to bottom up approaches.

Some important new work, outlined in two papers by Gerbig et al. (Gerbig *et al.* 2003a; Gerbig *et al.* 2003b), show a proof of concept on how to characterize land surface CO_2 fluxes at the continental scale, breaching the divide between top down and bottom up approaches. The two papers by Gerbig et al. describe an aircraft campaign, CO_2 Budget and Rectification Airborne (COBRA), that acquired CO_2 concentration measurements at high temporal and spatial resolution across the United States. These measurements were then used with a lagrangian tracer transport model and a simple plant physiology model to invert the observed CO_2 concentrations and obtain large-scale fluxes of CO_2 at a surface resolution of approximately 18 x 28 km across the northern and southern tier of the United States. Aircraft observations of CO_2 flux have been used in the past to estimate regional scale fluxes (Desjardins *et al.* 1995), however the Gerbig et al. study is the most ambitious evaluation to date of how to accurately characterize surface fluxes of CO_2 over very large areas. Among the many recommendations of the study was the development of appropriate analysis frameworks for high temporal and spatial variations in CO_2 flux which they found contribute most of the signal needed to interpret atmospheric CO_2 concentration variability at the landscape scale. This is of particular relevance considering the advent of the North American Carbon Program (NACP) (Wofsy and Harriss 2002) which will provide much higher resolution CO_2 flux and concentration data, both spatially and temporally, across the United States and Canada than has previously been available.

In this chapter, the effect of the representation of landscape heterogeneity on regional estimates of CO_2 flux is examined. It is clear from the aforementioned studies that the spatial and temporal representation of the land surface plays an important role in the interpretation of CO_2 concentration measurements and the development of regional estimates of CO_2 flux. Heterogeneity of the land surface is both scale dependent spatially as well as seasonally dependent. Using a coupled land surface-surface atmosphere model at the regional scale, simulations are conducted with two representations of the land surface. High heterogeneity simulations are conducted with a land surface increment of 1 km, lower heterogeneity simulations are conducted with a land surface increment of 8 km. Six simulations covering a range of landscape conditions, from late spring to high summer, were performed to answer the following questions:

1. What are the consequences of disregarding landscape heterogeneity on regional estimates of CO_2 flux?
2. Do regional estimates of latent and sensible heat flux respond similarly to loss of landscape heterogeneity?
3. Are there seasonal differences that need to be considered when assessing the effects of landscape heterogeneity on regional estimates of land surface fluxes of heat, water and carbon?

The work presented in this chapter contributes to the interpretation of the measurement and modeling of CO_2 at a variety of scales. It follows on from the previous two chapters where the sensitivity of the land surface model used in this study was evaluated and the uncertainty of fluxes due to the parameterization of the land surface model were evaluated. Here, however, we consider the uncertainty in flux estimates associated with the spatial representation of the land surface.

5.2 Background

Interactions between the land surface and the atmosphere have a notable effect on climate and weather patterns. This has been recognized and studied for some time now with respect to global climate and weather patterns (since at least (Charney 1975)) but vegetation-atmosphere interactions at regional to local scales have, until recently, been less well understood (Pielke *et al.* 1998; Pitman *et al.* 1999). The land surface and the atmosphere interact primarily through exchanges of energy in the forms of radiation, heat, moisture and momentum. The partitioning of incoming and outgoing energy into these forms and it's spatial variability have important implications for the interaction between the land surface and the atmosphere, for example how much moisture is available for precipitation, how much energy is available for boundary layer growth and how atmospheric circulations will develop. The interactions between vegetation and the atmosphere are important to studies of regional and global carbon balance because they regulate the conditions for photosynthesis, respiration and transport.

The presence, distribution and status of vegetation on the land surface have a large impact on how the available energy at the surface is partitioned into LE, H and G. For example, the presence or absence of vegetation on the land surface affects how much of the ground surface is directly exposed to incoming radiation. When more soil is exposed more energy can be partitioned into ground heat flux. However, the exposure of the soil to more direct radiation also leads to greater evaporation from the soil surface. As vegetation cover and leaf area increase, the albedo of the surface decreases. This leads to more absorption of radiation by the vegetation and increased latent heat flux in moist conditions, as there is more surface area for evapotranspiration (Stohlgren *et al.* 1998). The seasonal cycle of vegetation (phenology) is also important; when leaves are not present, more ground surface is exposed, the surface roughness is altered and sensible heat flux dominates latent heat flux, which can lead to a deeper boundary layer (Cotton and Pielke Sr. 1995; Pielke *et al.* 1998). The presence or absence of vegetation also influences the amount and type of litter (detritus and debris from plants) on the surface, which affects the absorption and emission of radiation by the surface.

At the regional scale the spatial composition of elements on the landscape affects land surface-atmosphere interactions. When the land surface is patchy (homogenous clusters of similar surface types and conditions), processes that generate gradients in temperature and pressure between adjacent patches can initiate local and mesoscale flow, redistributing energy through conduction and convection. For example, variability in surface albedo (Pielke *et al.* 1993), conductance of vegetation (Avissar and Pielke 1991), soil moisture (Segal *et al.* 1988) and land use (Chase *et al.* 1999; Pielke Sr. *et al.* 1999) are all conditions that have been shown to be capable of initiating this type of atmospheric flow. When the land surface is more homogeneous, the predominant characteristics of the area determine the partitioning of energy and the uniformity of the surface precludes gradient generated flow. Heterogeneous landscapes, where surface characteristics such as vegetation type, cover and water status are irregularly distributed, are more complex and more common. Unfortunately, they are also the least well understood with respect to their interaction with the atmosphere, in part because measurements over heterogeneous land surfaces are harder to interpret and in part because heterogeneous systems are difficult to model. Recent work has suggested that land surface heterogeneity at relatively small length scales (<5 km) plays an active role in forcing mesoscale flows in regional atmospheric models (Weaver and Avissar 2001; Roy and Avissar 2002; Roy *et al.* 2003).

The land surface is typically represented in two different ways in coupled land-surface atmosphere models. Lumped land surface models (such as SiB2; (Sellers *et al.* 1996b) and LEAF-2; (Walko *et al.* 2000)) use grid level parameters as input and produce grid level results and rely as much as possible on scale-invariant relationships. To represent more heterogeneity, smaller grid scales must be used or they must be tiled within the grid increment of the atmospheric model. Distributed land surface models account for subgrid scale variability through statistics or areal weighting.

Sellers et al. (Sellers *et al.* 1995b; Sellers *et al.* 1997b) examined the effects of land surface heterogeneity (soil moisture, vegetation cover and topography) on fluxes of sensible and latent heat modeled with SiB for the FIFE site in Kansas. Utilizing a variety of aggregation experiments, they found that the gentle topography in the FIFE area scaled linearly and thus had relatively little effect on calculated heat fluxes. The effects of vegetation cover on the heat fluxes also scaled linearly, partly because the land cover class was similar across their domain and partly because of the linear relationship between LE and f_{PAR} in SiB. They did find, however, that the heterogeneity effects of soil moisture were non-linear on the heat fluxes. The effects were most pronounced with moderately wet soil conditions, but as the land surface dried out soil moisture across the landscape became more self-similar and the effects became negligible. Other research suggests that the kind of scale invariance shown in the Sellers study does not hold for more heterogeneous land surfaces, those in which the vegetation cover is patchy, in which the land cover types are more complex or where the landscape is rapidly changing (Moran *et al.* 1997; Kimball *et al.* 1999; Chen *et al.* 2003; Yates *et al.* 2003). Many of these studies have been done without an interactive atmosphere, but as Margulis and Entekhabi (Margulis and Entekhabi 2001b) have shown, in coupled models sensitivities are often amplified. When Lu and Shuttleworth (Lu and Shuttleworth 2002) incorporated NDVI into a coupled simulation of the central United States they found the heterogeneity of the NDVI, along with diverse land cover types, altered the regional climate.

There is a dense body of literature on scaling (the preservation of some property of the landscape when going from finer to larger observational scales) and aggregation (what is lost when coarser observational scales are substituted for finer scales) of surface fluxes of sensible and latent heat (Avissar 1995; Wood 1995; Friedl 1996; Hu and Islam 1997; Moran *et al.* 1997; Kustas and Jackson 1999 - among others). Considerably less work has been done with respect to understanding how the flux and concentration of CO_2 in heterogeneous landscapes varies with scale of representation, though we might expect there to be differences. The differences would likely stem from the degree of coupling between the net ecosystem exchange of carbon and fluxes of latent and sensible heat as well as non-linearities within the coupling. Transpiration in plants is coupled to photosynthesis through a shared dependence on stomatal conductance. Stomata regulate the flow of water and carbon dioxide into and out of leaves. Whole plant water status is maintained through transpiration; when soil water availability is low or when atmospheric humidity is low stomata close, reducing conductance of water vapor to negligible levels in order to preserve a functional plant water status (Jones 1992).

Conversely, when concentrations of CO_2 in the leaves of plants drop below the level required for net photosynthetic uptake (i.e. the CO_2 compensation point which is ~130×10^{-6}mole fraction for C_4 plants, ~230×10^{-6} for C_3 plants), stomata *open* to allow CO_2 to diffuse into the leaf and bring it back above the compensation point (Jones 1992). The drop in intercellular CO_2 concentration is due to fixation. Plants are thought to balance the competing processes of water loss and CO_2 gain in order to optimize photosynthesis and/or growth. When water isn't limiting, transpiration facilitates the uptake of CO_2 because stomata are open and CO_2 can diffuse in freely. The relationship between CO_2 flux and transpiration thus serves as an added constraint on the surface energy balance as well as on CO_2 flux into and out of the canopy. In land surface-atmosphere modeling this relationship provides a critical link between biological processes and atmospheric processes at many scales because it is quantifiable and can be tested with observations at multiple scales.

The net ecosystem exchange (NEE) of carbon in an ecosystem consists not only of the flux of CO_2 into and out of the canopy, but also heterotrophic respiration. Heterotrophic respiration of carbon from the soil is the result of the decomposition of plant detritus (dead roots, litter) by microfauna, bacteria and fungi (Schlesinger 1997). Rates of decomposition in nature vary with temperature, moisture and the structure and chemical composition of the detritus. In SiB2.5, ecosystem respiration is dependant on temperature, moisture, soil texture and the annual uptake of CO_2 by the canopy (Raich and Nadelhoffer 1989; Denning *et al.* 1996a). The respiration rate is controlled by a simple exponential relationship for temperature (commonly referred to as Q_{10} and set to 2.4 in SiB2) and by three parameters related to soil texture. Thus respiration is coupled to both latent heat flux, via the effects of the vegetation on soil moisture and direct soil evaporation, and sensible heat flux via changing soil temperature.

There are other indirect effects that alter the net ecosystem exchange of carbon from the surface, such as boundary layer growth from sensible heat flux affecting the concentration of CO_2 above the canopy level. Furthermore, given the interrelationships and non-linearities that exist between latent and sensible heat flux and net ecosystem exchange, we might expect that different expressions of land surface heterogeneity will affect regional scale patterns in CO_2 exchange and CO_2 concentration differently than for either latent or sensible heat flux.

5.3 Methods

5.3.1 *The Regional Atmospheric Modeling System (RAMS)*

The Regional Atmospheric Modeling System (RAMS), developed at Colorado State University, is a flexible, comprehensive mesoscale meteorological modeling system (Pielke *et al.* 1992; Cotton *et al.* 2003). RAMS is a three-dimensional, non-hydrostatic model that solves the equations for atmospheric motion, thermodynamics, mass and water mixing ratio for a region of the atmosphere in a terrain following coordinate system. The standard version of RAMS is also coupled to land surface and hydrological sub-models that provide surface fluxes of heat, moisture and upwelling radiation (Walko *et al.* 2000). RAMS can be run for a variety of scales, from high-resolution large eddy simulations to coarse resolution hemispheric and even global simulations. A particularly advantageous feature of RAMS that is employed in this analysis is its nested grid capability (Walko *et al.* 1995). Nested grids are telescoping grids of coarse to fine computation and spatial resolution. They are computationally efficient because they allow a large region to be simulated while concentrating computational time on specific areas of interest. Coarser grids are updated in space and time with the progressively finer grids. Grids can be nested horizontally or vertically and can be stationary or move, as with a thunderstorm. For a thorough review of the mechanics of RAMS, the variety of applications it has been used for, and additional references see Pielke et al. (Pielke *et al.* 1992) and Cotton et al. (Cotton *et al.* 2003).

This analysis utilizes RAMS version 4.2. There are a variety of ways to configure RAMS. The possible options define not only such things as grid sizes and time steps, but also which schemes and parameterizations to use for turbulence, radiative transfer and lateral boundary conditions among others. The grid configuration for this analysis consists of 3 nested grids: a coarse grid (1) with 16 km horizontal spacing, extending 640 x 640 km, a secondary grid (2) with 4 km horizontal grid spacing extending 152 x 152 km, and an inner grid (3) with 1 km horizontal grid spacing extending 38 x 38 km. All three grids are centered on the WLEF tall tower and stretch to the top of the domain in the vertical (Figure 5.1). The time step for grid 1 is 45s, for grid 2 is 15s and for grid 3 is 5s. There are 45 variably spaced vertical levels with the finest levels nearer to the surface. The lowest layer of the domain is from 0 to 20m, the top is at 8341m. No vertical layer is greater than 800m depth. Nudging of the lateral boundary conditions towards NCEP reanalysis data (described below) was performed on the five outermost grid cells of grid 1. Nudging from the top of the domain was set to influence the atmosphere from 5500m to the top of the domain. Radiative lateral boundary conditions were used (Klemp and Wilhelmson 1978). There is no nudging in the interior of the domain. Recent research has shown that when modeling heterogeneous land surfaces, central nudging of the modeling domain both diminishes mesoscale circulations and increases the influence of the sub-grid turbulence parameterization, both undesirable effects (Weaver *et al.* 2002). Lateral and top nudging were found to have much less of an effect. The Harrington two-stream radiation parameterization (Harrington 1997) was used with the radiation updated every fifteen minutes. In a previous study (Nicholls *et al.* 2004) it was found that the model created anomalous fog when the microphysics parameterization was used at the WLEF site. Because of the excessive cooling and fog the microphysics parameterization was turned off and clear days as determined by GOES satellite imagery and surface radiation measurements were chosen to simulate. The Smagorinsky horizontal deformation scheme for turbulence (Smagorinsky 1963) coupled with the Mellor-Yamada (Mellor and Yamada 1982) vertical scheme employing prognostic turbulent kinetic energy was found to perform well at this location and for these simulation scales (Nicholls *et al.* 2004) and was used here.

Figure 5.1 Grid 1 (blue), Grid 2 (green) and Grid 3 (red).

5.3.2 *The Simple Biosphere Model(SiB)*

The Simple Biosphere Model (SiB2, SiB2.5) is described in Chapter 2 and in Sellers et al. (Sellers *et al.* 1996b) and Baker et al. (Baker *et al.* 2003). The version of the SiB model coupled to RAMS is SiB2c, which includes storage of carbon in the canopy airspace but not storage of heat and water. All other properties of SiB2 and SiB2.5 are the same.

5.3.3 *The coupled model (SiB-RAMS)*

The two models, SiB2c and RAMS are interactively coupled. RAMS provides to SiB2c atmospheric boundary layer values of air temperature, pressure, vapor pressure, wind velocity, direct and diffuse short-wave and long-wave radiation and CO_2 partial pressure. SiB2c provides to RAMS surface layer fluxes of heat, moisture, momentum, upwelling radiation and CO_2. Interaction between the models occurs at the lowest level above the land surface in RAMS every 60 seconds.

5.3.4 *Time invariant soil parameters*

Percent sand and percent clay for the study domain were derived from the State Soil Geographic Data Base (STATSGO) (Staff 1994). STATSGO is a high-resolution national soils database developed by the United States Department of Agriculture for regional studies. Detailed descriptions of the data collected, collection procedures and quality assessments are reported in the STATSGO metadata and users guide (Staff 1994). Percent clay by weight is given for each component for each layer in a soil mapping unit in the STATSGO database. However, percent sand and silt need to be derived from ancillary data; in this case particle size distributions. Silt was defined as the percent material by weight passing through a sieve with openings of 0.075mm minus the clay fraction. Sand was defined as the percent of material by weight passing through a sieve with openings not less than 0.075mm and not greater than 2.0mm and was calculated as the remainder of 100 – %silt – %clay. These percentages were calculated for each component for each layer in each soil-mapping unit. A weighted average was then calculated over the total depth of the profile for each component. Finally, average percent sand, silt and clay were calculated for the entire soil mapping unit by using the weighted contribution of each soil component to the soil mapping unit. The data was then gridded to a resolution of 1 km from its original polygon topology. Given that the smallest linear delineation in STATSGO is 1.25 km, this resolution was deemed reasonable. Soil hydraulic and thermal parameters were then calculated from the percent of sand and clay in the soil as described in Chapter 2.

5.3.5 *Time invariant vegetation parameters*

The time invariant vegetation biophysical parameters (27 in all) are based on values recorded in the literature and assigned to a particular biome via look-up tables (Sellers *et al.* 1996a). The 1 km vegetation classification created by Hansen et al.. (Hansen *et al.* 2000) was used to assign these biophysical parameters across the modeling domain. It was necessary to remap the Hansen dataset to SiB2 vegetation classes in order to assign the properties. Table 5.1 shows how the biome classes were remapped from the Hansen classes to the SiB2 classes for study domain. Only the biome classes present in the domain are presented.

Table 5.1. Biome classes from the Hansen *et al.* 2000 dataset were remapped into the SiB2 classes described in Sellers *et al.* 1996b for the mesoscale domain.

Class number and description	
Sellers et al. 1996b	*Hansen et al. 2000*
2 Tall broadleaf deciduous trees	**4** Deciduous broadleaf forest
3 Tall broadleaf and needleleaf trees	**5** Mixed forest **6** Woodland
4 Tall needleleaf trees	**1** Evergreen needleleaf forest
9 Broadleaf shrubs with bare soil	**8** Closed shrubland
11 No vegetation	**12** Bare ground **14** Urban and built up
12 Agriculture and grasslands	**10** Grassland **11** Cropland

5.3.6 *Time varying vegetation parameters*

Time varying vegetation parameters for SiB2 are calculated from an annual cycle of the Normalized Difference Vegetation Index. The 1 km global AVHRR NDVI dataset (Eidenshink and Faundeen 1994; Teillet *et al.* 2000) was used in this study. The development of the global 1km AVHRR data set was a project of the International Geosphere Biosphere Program (IGBP). This high-resolution global data has been calibrated, atmospherically corrected, and geometrically registered. Data from the project are limited to time periods between 1992 and 1996. The biome type map used in this study (Hansen *et al.* 2000) was developed from this data set.

The 1 km global NDVI product is available as 10-day maximum value composites. One year of the dataset, February 1995 through January 1996, was processed. Because January 1995 was not available, it was replaced with the January 1996 data. Monthly maximum value composites of NDVI from the 10-day maximum value composites were created. Monthly maximum value composites increase the chance of obtaining cloud free data. The registration between the soil maps, the vegetation classification and the NDVI data was then checked and adjusted so that water features in all layers were properly represented.

The equations and methodologies described in Sellers et al.. (Sellers *et al.* 1996a; Sellers *et al.* 1996b) and Los et al. (Los *et al.* 2000) and in Chapter 2 were used to calculate the time varying vegetation parameters from the NDVI data with the exception of one modification. The equations used to convert the NDVI data to f_{PAR} require the 2nd and 98th percentile value for the annual range in NDVI for each biome. These statistics must be recalculated for any new NDVI dataset to be used with SiB2 because different processing techniques and different resolutions of data can produce variable ranges of NDVI for the same locations. These percentiles were calculated using the regional subset of NDVI and the time varying parameters were then calculated.

5.3.7 *Topography*

The United States Geological Survey GTOPO30 topography product was used for this analysis. GTOPO30 is a global digital elevation model with 30-arc second (~1 km) horizontal spacing produced by the USGS. GTOPO30 was compiled from 8 different sources and the vertical accuracy at any given point is partially dependant on these sources. A thorough discussion of the sources of data and processing methods for GTOPO30 can be found at http://edcdaac.usgs.gov/gtopo30/gtopo30.asp. GTOPO30 is one of the standard data products released with RAMS.

5.3.8 *Atmospheric boundary conditions*

Atmospheric initial and boundary conditions of air temperature, geopotential height, relative humidity, zonal wind and meridional wind from the NCEP Reanalysis 1 data product were used. NCEP Reanalysis 1 data is provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/. The NCEP Reanalysis 1 data has a 2.5-degree horizontal resolution and 17 vertical levels. For this analysis, the four times daily (0Z, 6Z, 12Z, 18Z) pressure level product was used for the year 1997. The NCEP Reanalysis 1 data was prepared for use with standard RAMS isentropic processing techniques. No soundings or surface reanalysis data were used, as one of the main interests of this work is to analyze the effect of the surface parameterization on land-surface atmosphere interactions. CO_2 concentration was initialized uniformly at 360 ppm.

5.3.9 *Water Temperature*

The temperature of the water bodies in the simulations was initialized with Great Lakes buoy data from the National Oceanic and Atmospheric Administration for the first day of the simulation http://www.nodc.noaa.gov/BUOY/buoy.html. For April, May and early June the water temperature was initialized as 275.7 degrees Kelvin, for July it was initialized as 290 degrees Kelvin, corresponding to buoy temperatures at these times. Large water bodies have large thermal inertia and the lack of difference between April and early June is evidence of this. These temperatures are appropriate for the Great Lakes, but may be less so for the smaller lakes in the simulation since they are less deep and much smaller. However, it was the best available information.

5.3.10 *Time-varying soil parameters*

Soil temperature and moisture are very difficult to initialize in mesoscale models, yet they are critical to robust simulations of land surface fluxes. Attempts have been made with thermal and radar remote sensing to measure the temperature and moisture status of the land surface over large areas, however the results so far suggest that only the top few centimeters of the soil are measurable by the sensors. Applying point measurements from soil temperature and moisture profiles to large areas is also an option, however it is neither appropriate nor adequate since both vegetation cover and soil type vary independently across the land surface. In a model such as RAMS, soil temperature is often related to air temperature and then scaled with depth. Soil moisture is very often initialized homogeneously across a region or spun up with the model for a relatively short period of time. It has generally not been possible to spin up soil temperature and moisture for appropriate lengths of time at the appropriate resolution because of the computational costs, though research has shown that well resolved soil moisture and temperature is important at the mesoscale (Eastman *et al.* 1998).

For this analysis, maintaining the integrity of the spatial heterogeneity of the land surface was very important. For this reason we decided to produce spatially resolved soil temperature and moisture profiles using long-term model spin-up. An area of 100 by 100 km at 1 km resolution centered on the WLEF tall tower was sampled from the spatial datasets described above. Each grid point in this array was spun up using the standard version of SiB2 for 4 annual cycles with the observed weather at WLEF for 1997. The initial soil moisture and temperature were those for the standard WLEF SiB2 run described in Chapter 2. Because each point had different soil and vegetation characteristics, the results were spatially resolved soil temperature and moisture profiles for each month of the year 1997. A spatially resolved canopy temperature was also calculated in this process. While observed weather at WLEF may not be representative of the weather at every point in this 100 x 100 km grid, it seemed preferable to the use of interpolated weather from much coarser scale reanalysis products.

The 100 x 100 km spun up soil moisture and temperature profiles and canopy temperature only accounted for the center portion of the whole 640 x 640 km domain. Since it was not feasible to spin up the whole domain and since the majority of the analysis in this chapter focuses on the 38 x 38 km inner grid, an approximation was made for the rest of the domain. Canopy temperature and soil surface temperature were averaged by biome from the spun up data and applied by biome to the rest of the domain. The soil temperature and moisture profiles were more complicated. Each grid cell in the entire domain was reclassified from the percentage of sand and clay to the appropriate

USDA soil texture class. The grid cells from the spin up were then sorted by biome and by soil texture class and an average soil temperature and soil moisture profile calculated. These average profiles were then applied by biome and texture class to the area outside of the spun up 100 x 100 km grid.

The calculation of spatially resolved soil respiration was incorporated into the spin up of the 100 x 100 km grid subset. Soil respiration in SiB2 is a function of the annual sum of carbon assimilated by the canopy, soil temperature and soil moisture expressed in the model through a term called respfactor (see Chapter 2 for a more thorough description). A different respfactor is assigned to 5 of the 7 soil temperature levels (the deepest levels below the roots do not respire carbon in SiB2). For the domain outside of the 100 x 100 km subset, respfactor for each soil level was averaged for each biome within the 100 x 100 km subset and then applied by biome to the rest of the domain.

5.3.11 *Aggregation procedures*

To simulate the loss of land surface heterogeneity, aggregation of land surface properties from the 1 km 'fine scale' to an 8 km 'coarse scale' was necessary. For soil percent sand and clay, a block mean process was used. Starting in the upper left corner of the soil maps, an 8 x 8 km region was averaged for percent sand and percent clay. Pixels that were comprised of water (and therefore had no soil information) were left out of the averaging process. Each 1 km grid cell in the 8 x 8 km block was then assigned this average value (Figure 5.2). Disaggregating the 8 km aggregated data back into 1 km grid cell sizes, reduced the land surface heterogeneity while maintaining the same land surface grid increments between simulations. For the biome type map, again starting in the upper left corner, a dominant class was determined for an 8 x 8 km region. The dominant class was then assigned to each 1 km grid cell in the 8 x 8 km region (Figure 5.3). For the NDVI data the process was broken down into two parts. For a subset of the domain centered on WLEF, 120 km x 110 km, the individual bands used to calculate NDVI were averaged to 8 km. NDVI was then calculated from the averaged bands. For the rest of the domain the NDVI itself was averaged to 8 km (Figure 5.4). The two techniques produce very similar results, however to be precise the individual bands should be averaged (Hall *et al.* 1992). The reason this was not done for the entire domain is because a faulty water mask was applied to the individual bands for the months of February, March, April and May. The faulty water mask affected 2.4% of the pixels in this subset and these pixels were carefully corrected with the nearest neighboring land pixel. It would have been a significant amount of work to correct the entire domain for little additional benefit, since the focus of this analysis is on the 38 x 38 km inner grid.

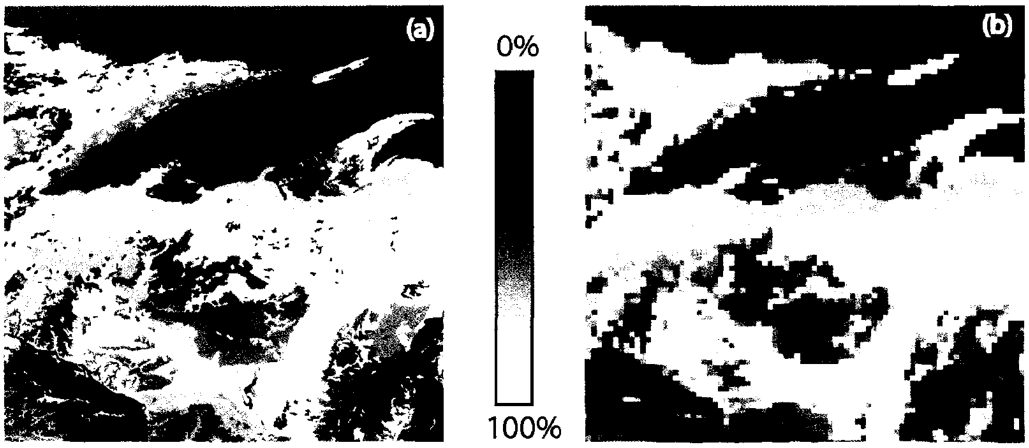


Figure 5.2 Percentage of sand in the soil a) 1 km b) 8 km.

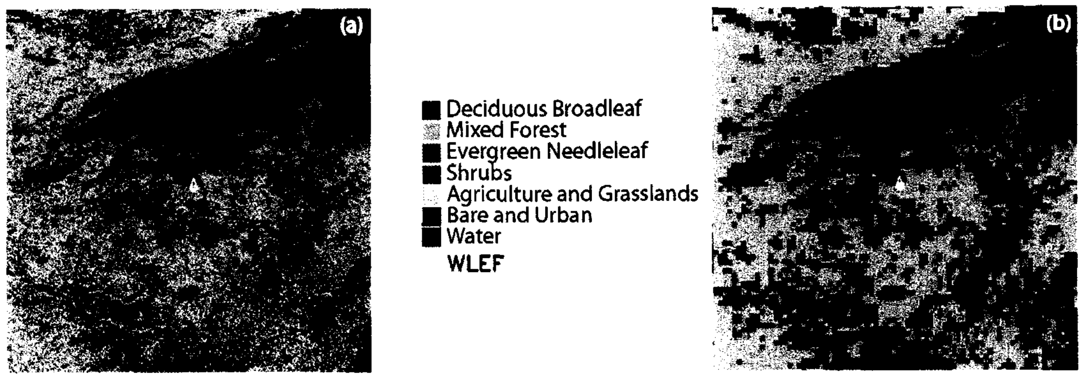


Figure 5.3 Biome classification map a) 1 km b) 8 km

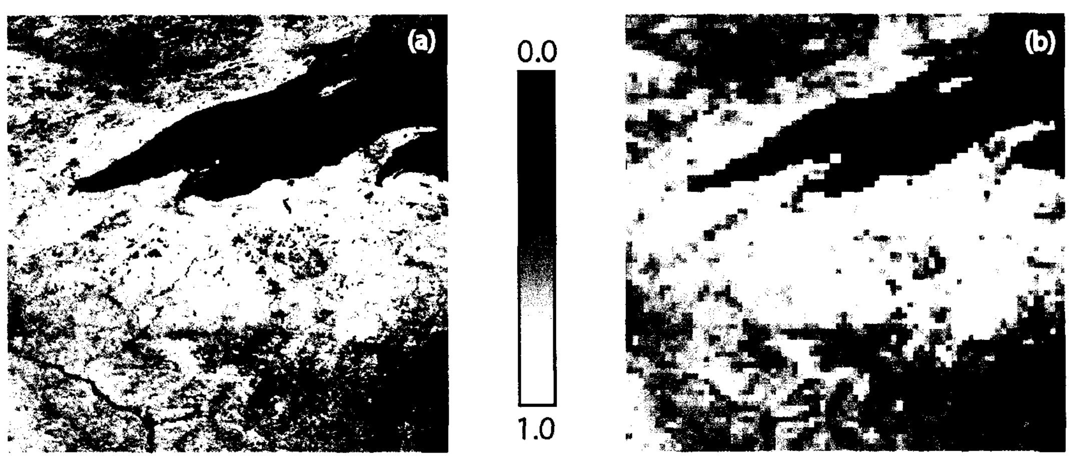


Figure 5.4 NDVI July 1997 a) 1 km b) 8 km.

5.3.12 *Experiment*

Three five-day periods, which were relatively free of clouds and had no measurable precipitation, were simulated: 12:00 GMT April 11 - 12:00 GMT April 16, 12:00 GMT May 30 - 12:00 GMT June 4, and 12:00 GMT July 26 - 12:00 GMT July 31, 1997. For each time period two simulations were run, one representing greater land surface heterogeneity (1 km land surface representation) and one representing reduced land surface heterogeneity (8 km land surface representation). In all simulations the actual horizontal grid increment for the finest grid was 1 km. Recent research investigating the effects of model configuration on mesoscale atmospheric circulations (utilizing the RAMS model) showed that there was a substantial difference in the character of circulations generated, both qualitatively and quantitatively, as the horizontal grid increment is increased above 1 km (Weaver *et al.* 2002). Weaver et al. found that coarser grid increments (even moving from 1 km to 2 km) resulted in broader and weaker updraft regions and diminished latent heat flux well into the boundary layer. Since the objective of this study was to evaluate the effects of the representation of land surface heterogeneity, not grid increment, maintaining the grid increment between simulations was considered the best approach.

Three time periods were selected to investigate because the different characteristics of the land surface during these periods reflect quite different stages of the annual carbon cycle in temperate ecosystems. Further, the sensitivity and uncertainty analyses of the previous chapters show quite different responses to model paramterization at these times. April, as discussed in the last chapter, was a time period of rapid surface change as the snow melted and the ground surface began to heat up. Figure 5.5 shows the meteorological conditions for the period just before, during and after the April simulation. Just preceding the beginning of the simulation, air temperatures are below freezing. During the period of the simulation, air temperatures progressively warm up, the relative humidity increases and windspeeds increase. The winds also change from north, northwest to west, southwest.

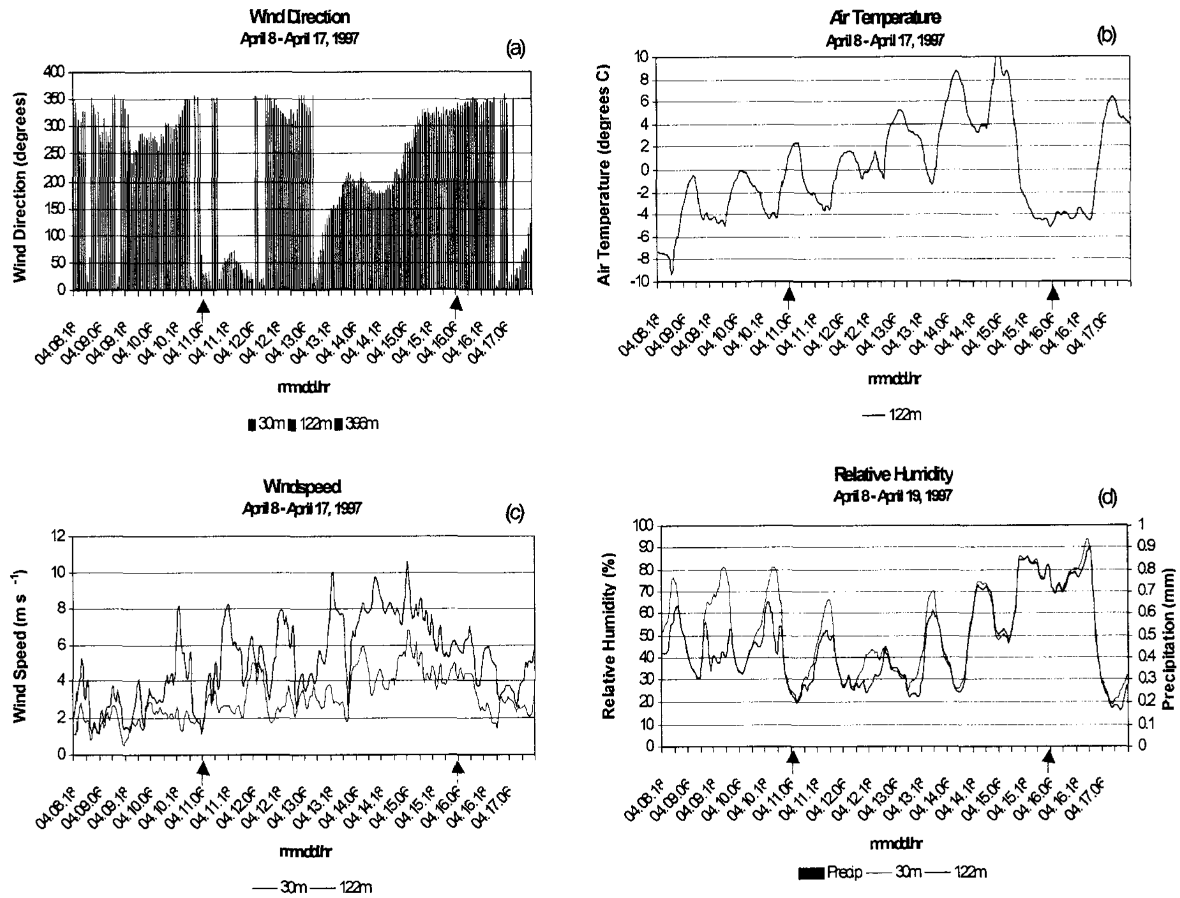


Figure 5.5 Observed weather conditions at the WLEF tower April 8 – April 17. Red arrows indicate the beginning and end of the simulation period. **a)** wind direction, **b)** air temperature, **c)** wind speed and **d)** relative humidity and precipitation.

Late May, early June is a time period of great sensitivity in the model since in 1997 leaf out of the deciduous trees took place on and around June 1st. Figure 5.6 shows the meteorological conditions for the period just before, during and after the late May, early June simulation. In the days preceding the simulation, conditions were humid, cool and there was a precipitation event. During the simulation period, winds were initially from the west but became more southeasterly during the middle of the simulation period before returning to westerly winds. Windspeeds were fairly consistent throughout the simulation period.

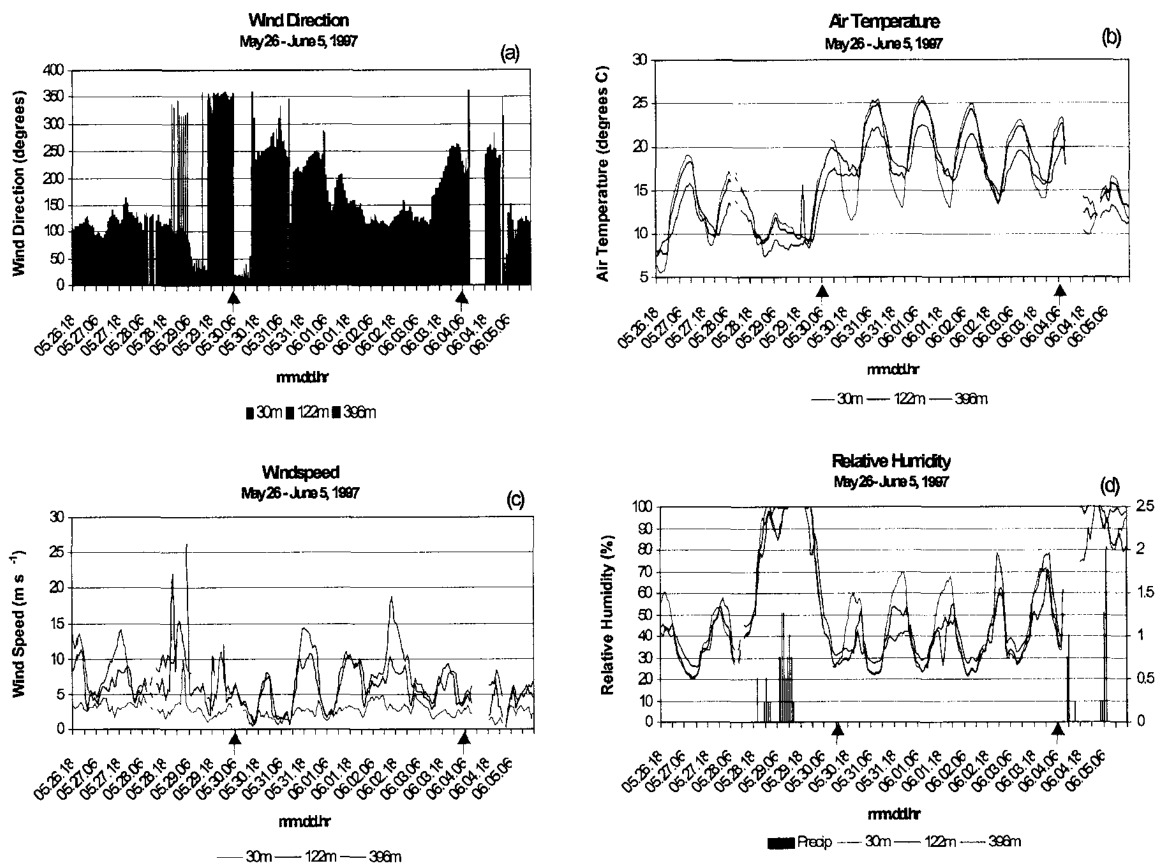


Figure 5.6 Observed weather conditions at the WLEF tower May 28 – June 5. Red arrows indicate the beginning and end of the simulation period. **a)** wind direction, **b)** air temperature, **c)** wind speed and **d)** relative humidity and precipitation.

July was a time period of little land surface change, and low sensitivity for the model, since the vegetation in the area of WLEF was at peak or near peak leaf area and the ground surface was effectively covered. Figure 5.7 shows the meteorological conditions for the period just before, during and after the late July simulation. Just before this simulation period, there was a significant rain event. Air temperatures were quite warm, but dropped as a cold front came through two days into the simulation period. Conditions were generally humid, particularly near the surface. Wind direction shifted from north, northwest in the beginning of the simulation period to west, southwest at the end of the simulation period. Windspeeds show an increase around the time the front came through but are otherwise consistently below 5 m s^{-1}.

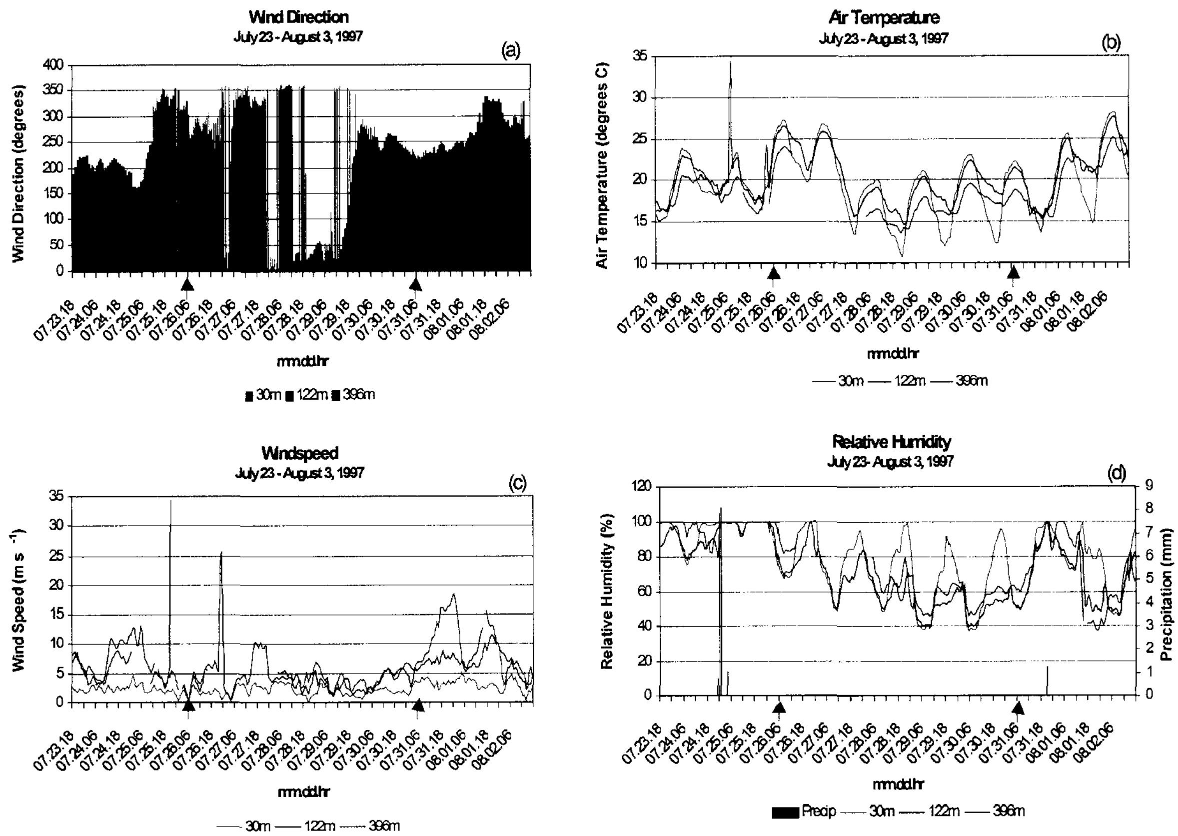


Figure 5.7 Observed weather conditions at the WLEF tower July 23 – August 3. Red arrows indicate the beginning and end of the simulation period. **a)** wind direction, **b)** air temperature, **c)** wind speed and **d)** relative humidity and precipitation.

5.4 Results

5.4.1 *Surface fluxes*

The following three figures show the ability of the SiB2-RAMS model to simulate surface fluxes of latent and sensible heat and the net ecosystem exchange of carbon at the tall tower site. Each figure shows the observations of each flux plotted with results from the high heterogeneity (1 km land surface) simulations. The modeled fluxes are averages of the four center grid points of the simulation centered on the WLEF tower site. Fluxes from the tall tower at WLEF represent a reasonably large portion of the landscape (see Chapter 2) so it seemed most appropriate to do this.

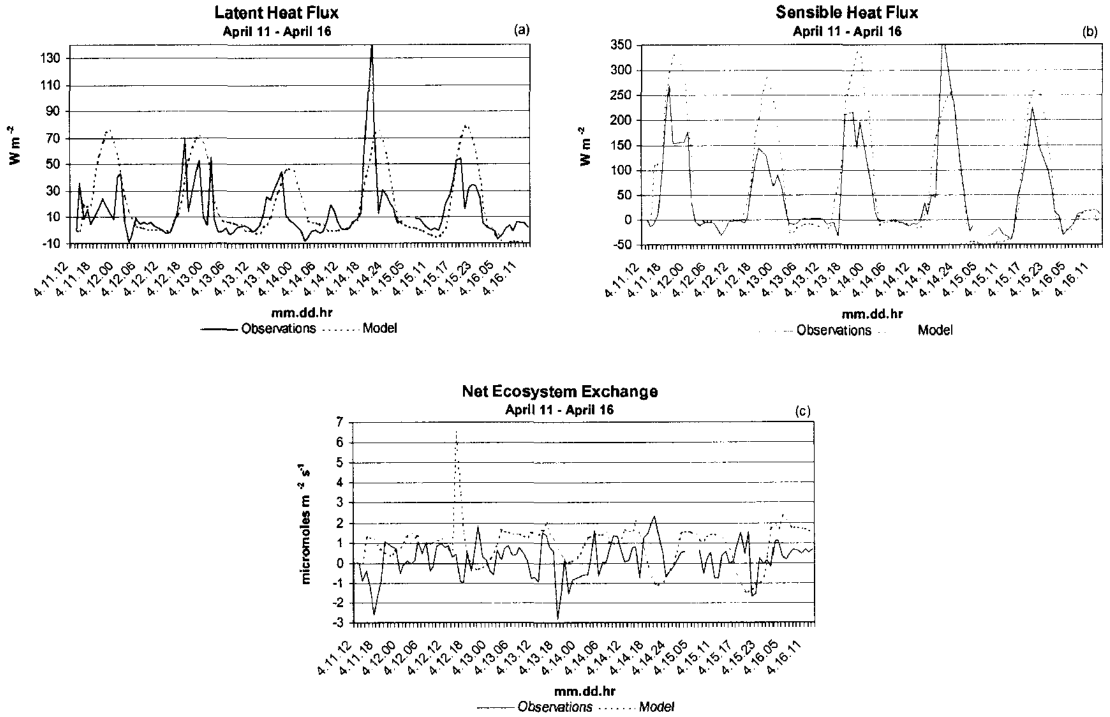


Figure 5.8 Observed (solid line) and modeled (dashed line) of a) latent heat flux b) sensible heat flux and c) net ecosystem exchange for April 11 thru 16th, 1997.

Figure 5.8 shows results from the mid-April simulations. The agreement between modeled fluxes and observations is reasonably good. For latent heat flux there is a tendency for the model to underestimate fluxes during the daytime. For sensible heat flux, the model initially overestimates the flux during the day but as the surface warms up, the agreement between the observations and the simulation improves. The magnitude of the net ecosystem exchange of carbon is small at this time of year and the model does a good job of capturing the small flux. There is an anomalous respiratory spike on April 12th. Examination of modeled surface temperatures, soil moisture and other variables don't explain this spike.

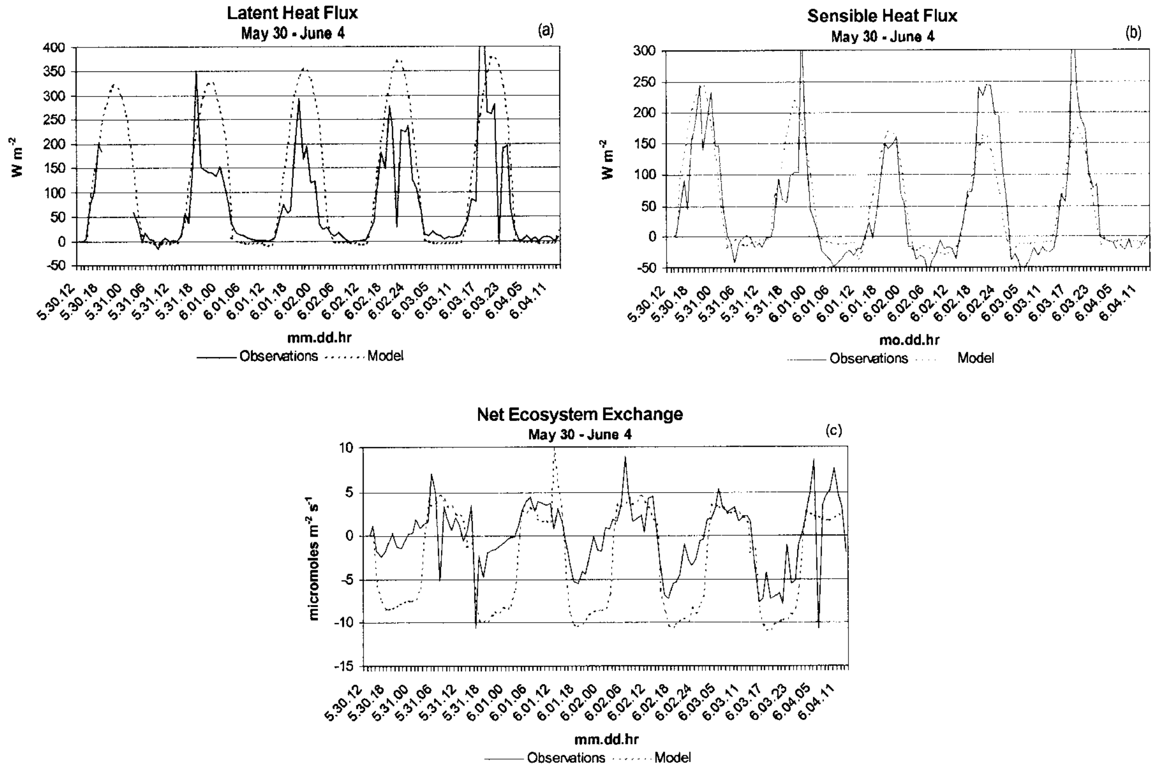


Figure 5.9 Observed (solid line) and modeled (dashed line) of a) latent heat flux b) sensible heat flux and c) net ecosystem exchange for May 30th thru June 4th, 1997.

Figure 5.9 shows results from the late May/early June simulations. The model overestimates fluxes of latent heat during the daytime, however agreement with the observations is better towards the end of the simulation period. For sensible heat flux, agreement between modeled and observed fluxes is quite good. Predicted net ecosystem exchange shows a similar pattern to the latent heat flux where the model overestimates net ecosystem exchange during the daytime at the beginning of the simulation period but agreement with the observations improves towards the end of the simulation. Leaves on deciduous trees began to unfurl at the WLEF site around June 1st. It is possible, as has been mentioned in previous chapters, that the leaf area is overestimated by the satellite data for this time period but that the overestimation is reduced as the leaves actually appear.

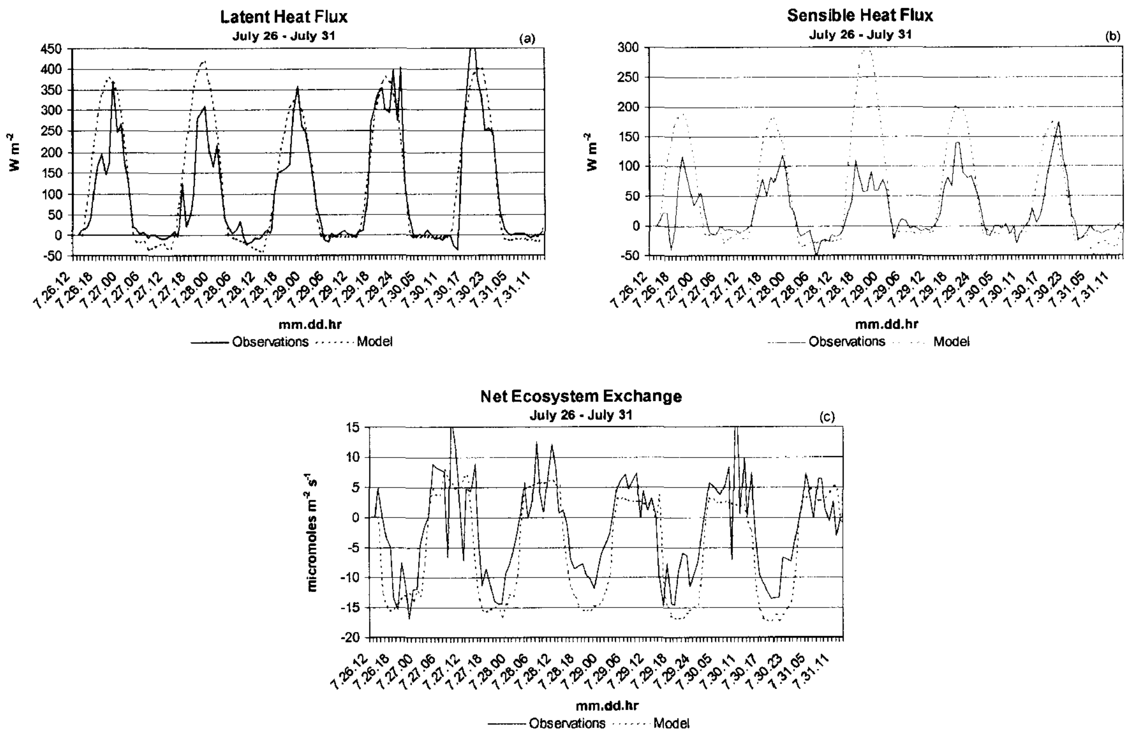


Figure 5.10 Observed (solid line) and modeled (dashed line) of a) latent heat flux b) sensible heat flux and c) net ecosystem exchange for July 26th thru July 31st, 1997.

Figure 5.10 shows results from the late July simulations. Modeled fluxes of latent heat and the net ecosystem exchange of carbon agree quite well with the observations during this time period. There is a slight overestimation of net ecosystem exchange during the daytime for the last three days of the simulation, but the fluxes appear to be within the observed uncertainty of the SiB model reported in the last chapter. Fluxes of sensible heat are overestimated in the daytime during this time period, a known problem with SiB in summer months (Baker *et al.* 2003).

Overall, agreement between the model and observations at the WLEF tall tower is reasonably good for all of the time periods simulated. The coupled model is able to capture the diurnal cycle well, though the variability seen in the observations at the hourly time scale is not captured. It is possible that for some days the incoming radiation from the model, which assumes clear skies, is an overestimate of the actual incoming radiation. Even though an attempt was made to pick clear days it is possible there were cirrus or patchy clouds at the site. Nonetheless, the coupled models are able to simulate land surface fluxes, the heterogeneity of which is the subject of the rest of the results.

5.4.2 *Spatial heterogeneity and grid mean fluxes*

Figures 5.11, 5.12 and 5.13 show examples of the effects of the representation of high and low land surface heterogeneity on surface fluxes of CO_2 for the three simulation periods. Each figure shows nighttime (midnight; simulation hour 91) and daytime (noon; simulation hour 103) flux maps for net ecosystem exchange for the innermost grid. In general there is a wider range in the magnitudes of the fluxes in the high heterogeneity simulations. Homogeneous patches of surface fluxes in the low heterogeneity land surface are maintained in the April simulations, however the large patches are less distinct in the later simulations, perhaps because of warmer surface and air temperatures leading to differential heating and cooling of the land surface and mixing with the atmosphere. The coupled model was parameterized with zero CO_2 flux over the water bodies. Water bodies do exchange heat fluxes with the atmosphere.

To examine the effects of the representation of land surface heterogeneity on regional predictions of net ecosystem exchange, latent and sensible heat flux, grid mean averages and cumulative grid mean sums of each flux were calculated through time for each simulation period for both the high and low heterogeneity simulations. The grid mean averages were calculated as a spatial average across the entire inner gird, for each simulation hour. Cumulative grid mean sums were calculated to evaluate whether differences in the high and low heterogeneity simulations result in significant drift in overall carbon, water or energy balance.

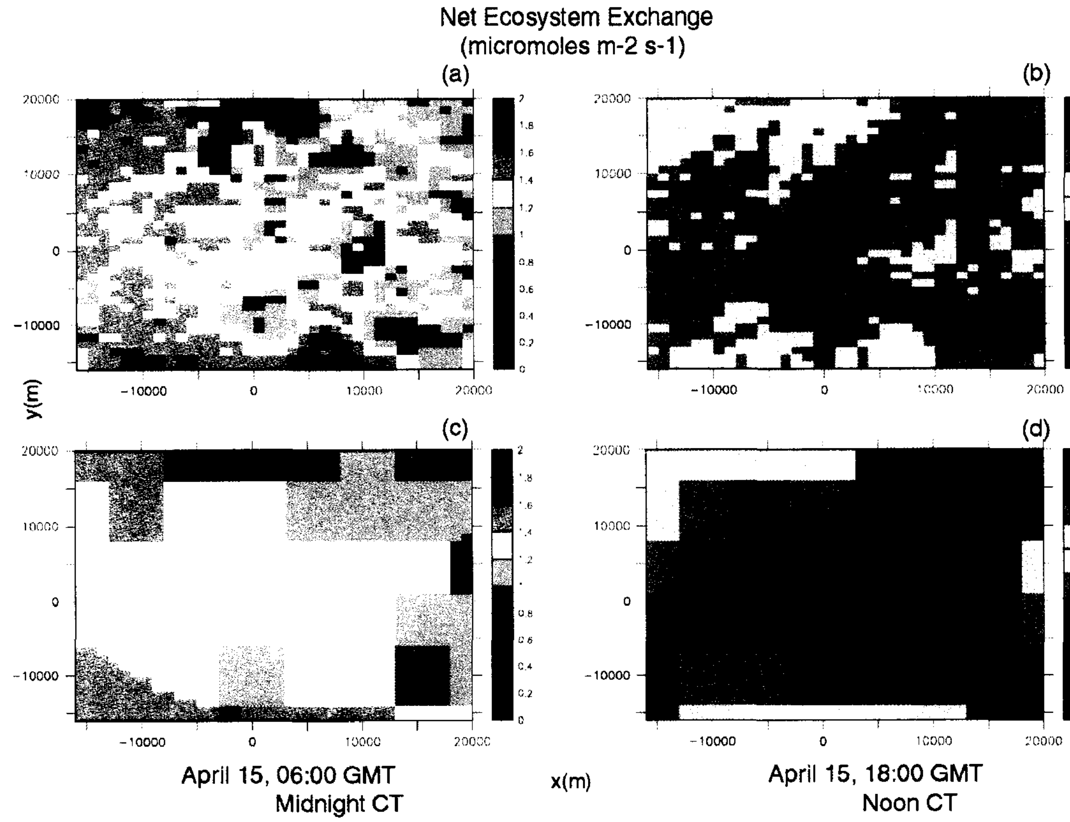


Figure 5.11 Net ecosystem exchange for two representations of land surface heterogeneity for two times of day. **a)** and **b)** represent the high heterogeneity simulations with a 1 km representation of heterogeneity. **c)** and **d)** represent the low heterogeneity simulations with an 8 km representation of heterogeneity. **a)** and **c)** show the net ecosystem exchange of carbon at midnight local time on April 15th, **b)** and **d)** show the net ecosystem exchange of carbon at noon local time on April 15th

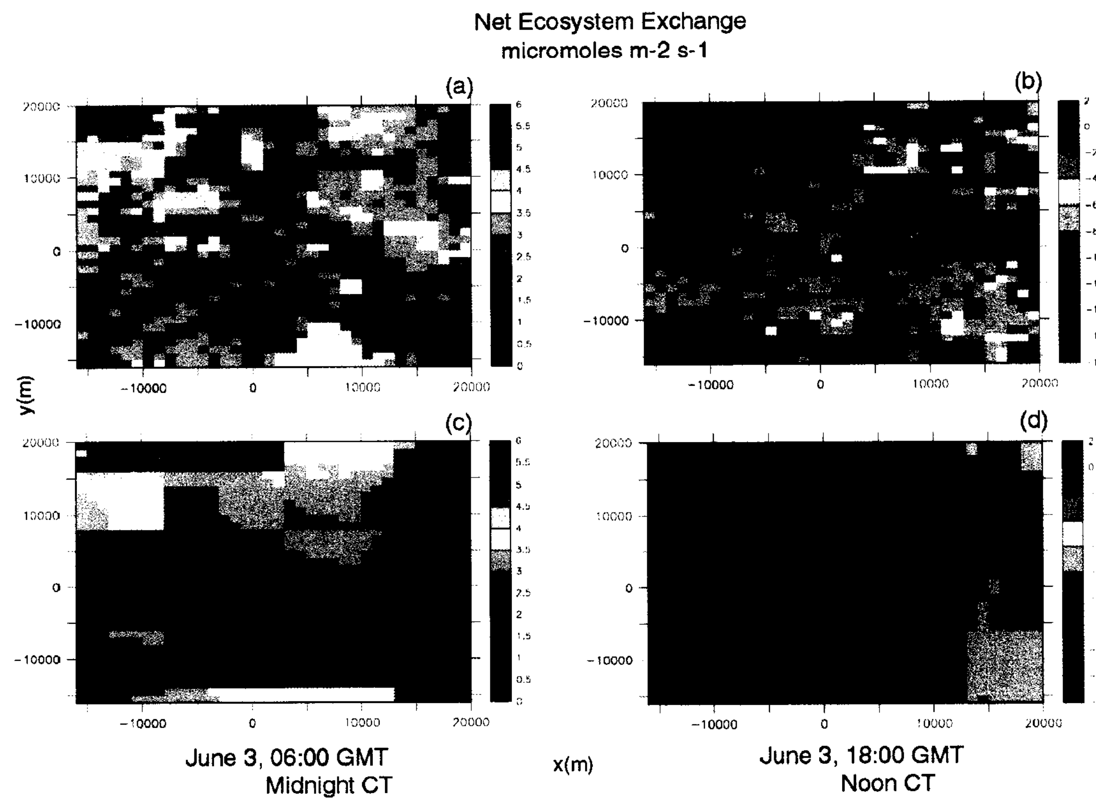


Figure 5.12 Same as Figure 5.11 but for June 3.

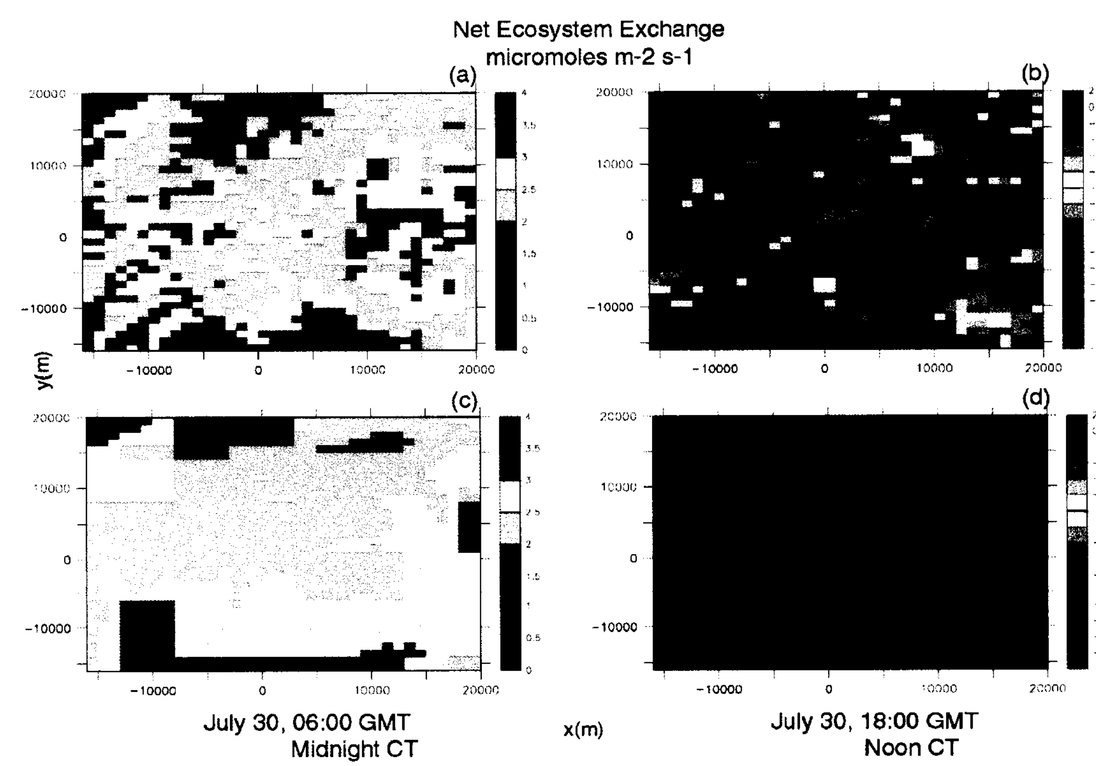


Figure 5.13 Same as Figure 5.11 but for July 30.

Figures 5.14, 5.15 and 5.16 show the results of the grid mean average fluxes through time and their cumulative sum. Figure 5.14 shows the results for net ecosystem exchange. For each time period, the high heterogeneity (1 km land surface) grid mean average estimated uptake during the day is slightly less than the low heterogeneity (8 km land surface) grid mean average. In April, where respiration exceeds assimilation, this results in a greater loss of carbon for the high heterogeneity simulation. In the late May/early June and July simulations it results in a smaller gain of carbon per meter squared. Table 5.2 gives the ratio of the cumulative grid mean average net ecosystem exchange at the last time step of the simulation for the low to high heterogeneity simulations. The results show a greater accumulation of carbon in the low heterogeneity simulations - 2% more for the July simulation and 7% more for the late May/early June simulation and a 3% greater loss of carbon for the April simulation.

Figure 5.15 shows the results for sensible heat flux. For each time period, the high heterogeneity grid mean average estimated sensible heat flux during the day is slightly less than the low heterogeneity grid mean average. This indicates that slightly more energy is lost to sensible heat flux in the low heterogeneity than in the high heterogeneity simulations. Table 5.2 shows the ratio of the low to high heterogeneity cumulative grid mean flux for each time period. The results show a greater flux of sensible heat out of the land surface for low heterogeneity simulations - 2% more for the April and late May/early June simulation and 5% more for the July simulation.

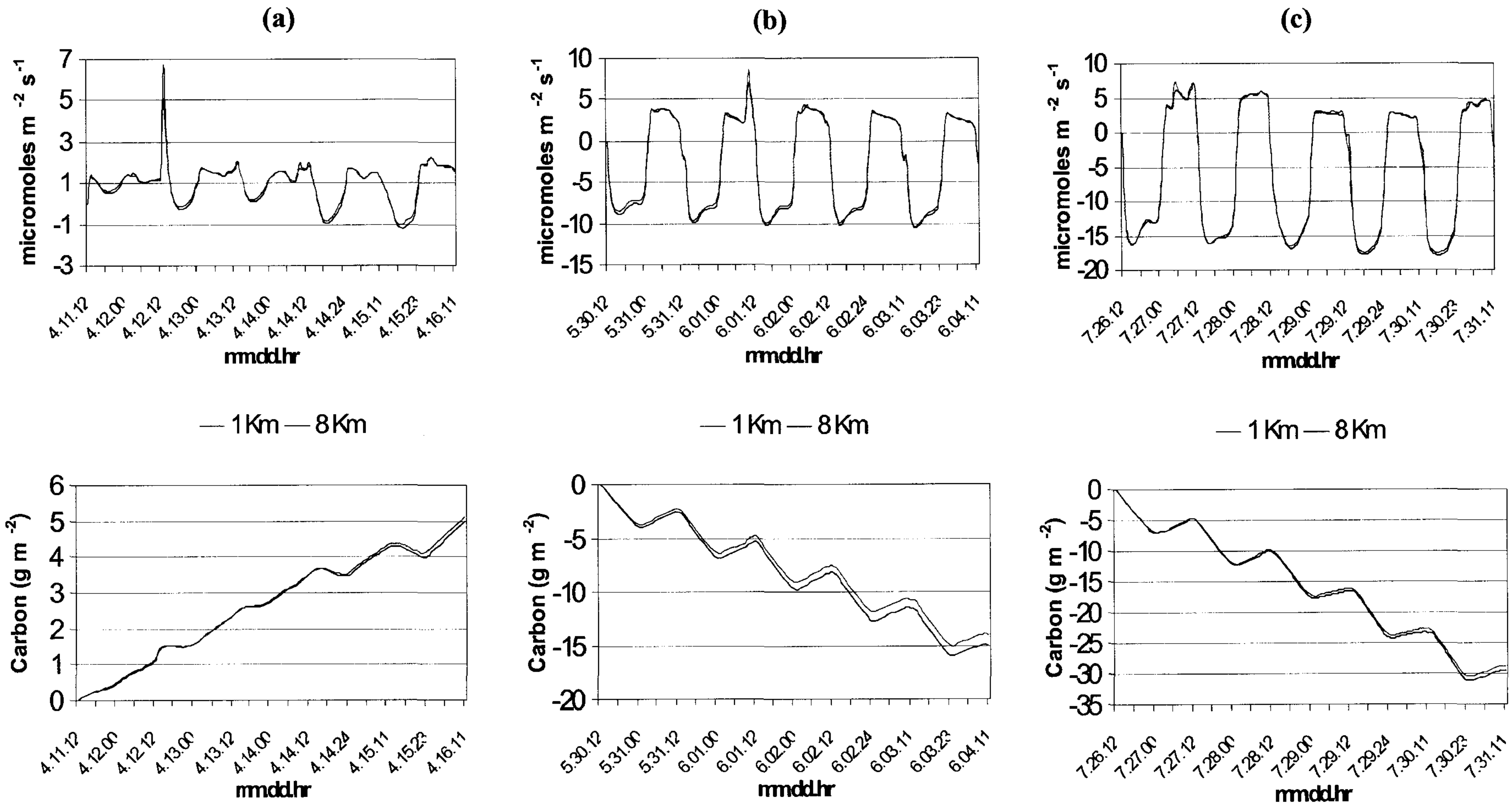


Figure 5.14 Grid mean averages (top row) and cumulative grid mean sums (bottom row) of net ecosystem exchange for the high (1 km) and low (8 km) heterogeneity simulations for a) April 11 – 16, b) May 30 – June 4 and c) July 26 – 31.

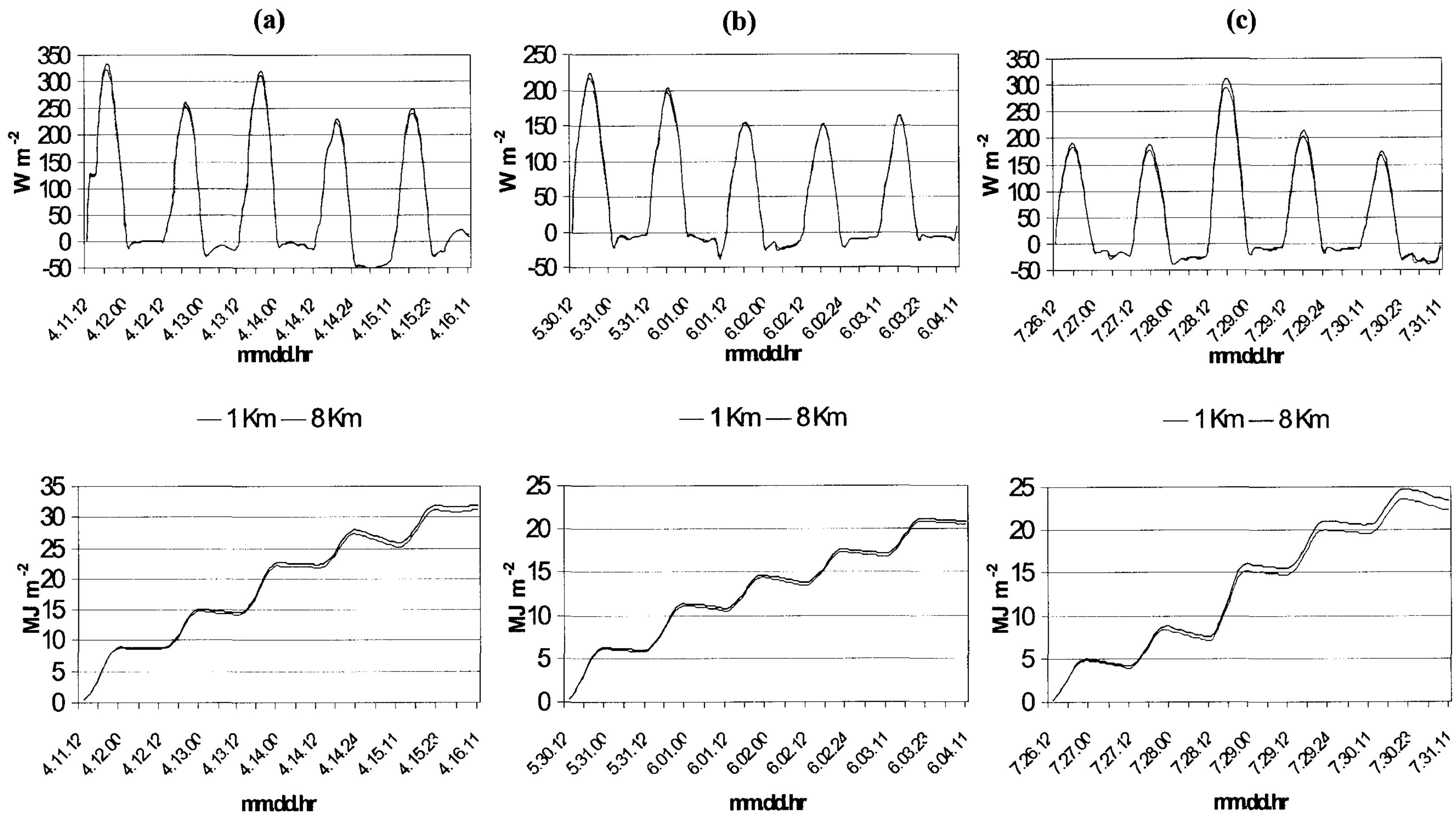
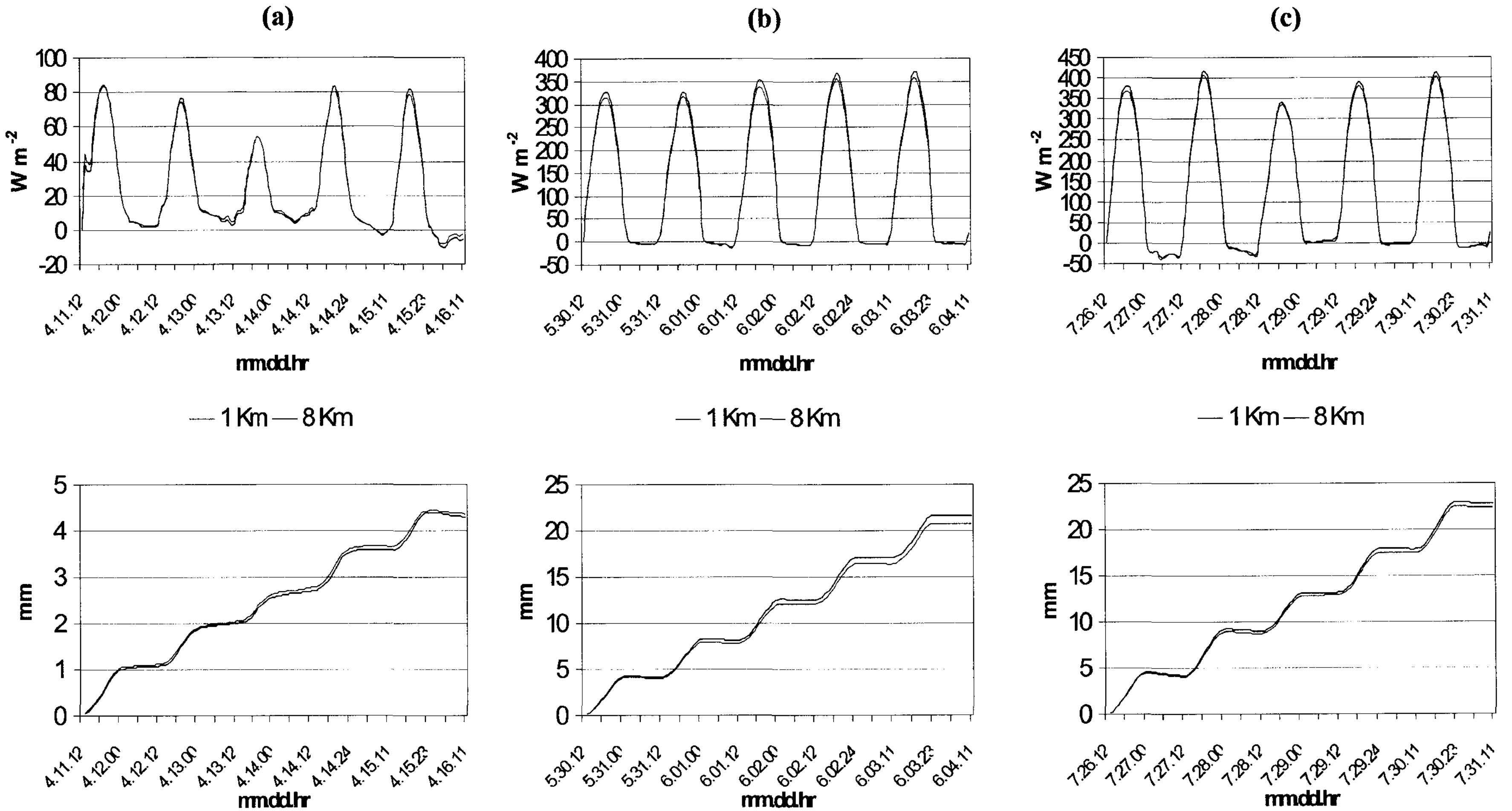


Figure 5.15 Grid mean averages (top row) and cumulative grid mean sums (bottom row) of sensible heat flux for the high (1 km) and low (8 km) heterogeneity simulations for **a**) April 11 – 16, b) May 30 – June 4 and c) July 26 – 31.



5.16 Grid mean averages (top row) and cumulative grid mean sums (bottom row) of water for the high (1 km) and low (8 km) heterogeneity simulations for a) April 11 – 16, b) May 30 – June 4 and c) July 26 – 31. Latent heat fluxes were converted to estimates of water using the latent heat of vaporization and the mean air temperature in each simulation period.

Figure 5.16 shows the results for latent heat flux. In April, the high heterogeneity simulations predict slightly more latent heat flux during the evening and slightly less during the day. The result is a slightly larger cumulative total latent heat flux for the high heterogeneity simulations. The late May/early June and July high heterogeneity simulations show a slightly lower prediction of latent heat flux during the daytime than the low heterogeneity simulations, resulting in slightly less predicted latent heat flux overall. Ratios of low to high heterogeneity cumulative grid mean flux for each time period are shown in Table 5.2. The results show a 4% greater flux of latent heat out of the land surface for the low heterogeneity simulations of late May/early June and 2% more from the July simulation, for the April simulation the high heterogeneity simulation shows a 2% greater flux.

Table 5.2 Ratios of the cumulative grid mean sum of high to low heterogeneity net ecosystem exchange (NEE), latent heat flux (LE) and sensible heat flux (H) for the three simulated time periods. Calculations including water features are outside of the parentheses, calculations without water features included are within parentheses.

Ratio of low to high heterogeneity			
	Water(*No Water*)		
	NEE	H	LE
April	0.97(*0.95*)	1.02(*1.00*)	0.98(*1.02*)
June	1.07(*1.05*)	1.02(*0.98*)	1.04(*1.01*)
July	1.02(*1.00*)	1.05(*1.02*)	1.02(*1.01*)

When dominant land surface types are used to characterize the landscape in coarse resolution simulations, such as the low heterogeneity simulation presented here, there is often a loss of diversity of land cover types. Table 5.3 lists the differences in land cover types between the high and low heterogeneity simulations described in this chapter. One of the landscape features that is often reduced in low spatial resolution simulations, and sometimes even eliminated, are water features. For the simulations presented here, there was a reduction in the representation of the water class of 3% in the low heterogeneity surface representation. The geographic configuration of the water was also different (see Figures 5.11-5.13). To understand the impact of the presence or absence of water on the grid mean surface fluxes, the cumulative grid mean flux for each simulation was calculated without the inclusion of water features. The results are presented in Table 5.2 as the ratio of the cumulative grid mean flux of the low heterogeneity simulations to hgih heterogeneity simulations. For the month of July the result was less of a difference between the two simulations for all of the fluxes, though there was still a 2% difference between the simulations for sensible heat flux. For the late May/early June simulation there was also a reduction in the differences between the two simulations, however there remained a 5% difference between simulations in net ecosystem exchange. For the April simulations, the difference in net ecosystem exchange increased a small amount from 3 to 5%. The difference in latent heat flux remained the same, however the direction of the difference changed. Without the water included, the high heterogeneity simulation in April showed less latent heat flux than the low heterogeneity simulation.

Table 5.3 Change in land surface classes from high to low heterogeneity land surface representation.

Class number and description		**Percent Representation Grid 3**	
		High	*Low*
2	Tall broadleaf deciduous trees	31	15
3	Tall broadleaf and needleleaf trees	59	81
4	Tall needleleaf trees	3	0
14	Water	7	4

Overall, the grid mean differences were not very large; though over time small differences would eventually add up. For example, after only 5 days the difference between the low and high heterogeneity simulations was 0.7 g C m^{-2}. However, over smaller spatial areas the differences are likely to be much greater. Figures 5.17 through 5.19 show the grid mean 5^{th} and 95^{th} percentile deviations between the high and low spatial heterogeneity simulations and the cumulative sum of those deviations for net ecosystem exchange, sensible heat flux and latent heat flux respectively. The deviations were calculated by subtracting the low heterogeneity simulations from the high heterogeneity simulations at each point (excluding all grid cells designated as water in either configuration), ranking them and calculating the 5^{th} and 95^{th} percentiles. The 5^{th} and 95^{th} percentiles were used here to represent the spatial uncertainty resulting from different representations of land surface heterogeneity. More than 10% of the grid cells have deviations and cumulative deviations that are greater than the percentile lines plotted.

Even though there was very little difference between the high and low heterogeneity simulations on average across the inner grid, Figure 5.17 shows that the differences at individual points can be quite large. As might be expected for net ecosystem exchange, as the land surface becomes more active and the fluxes larger, potential differences become greater. The differences are greatest during the daytime and result in potentially large differences in carbon gain/loss from the land surface, up to 10 g C m^{-2} after 5 days.

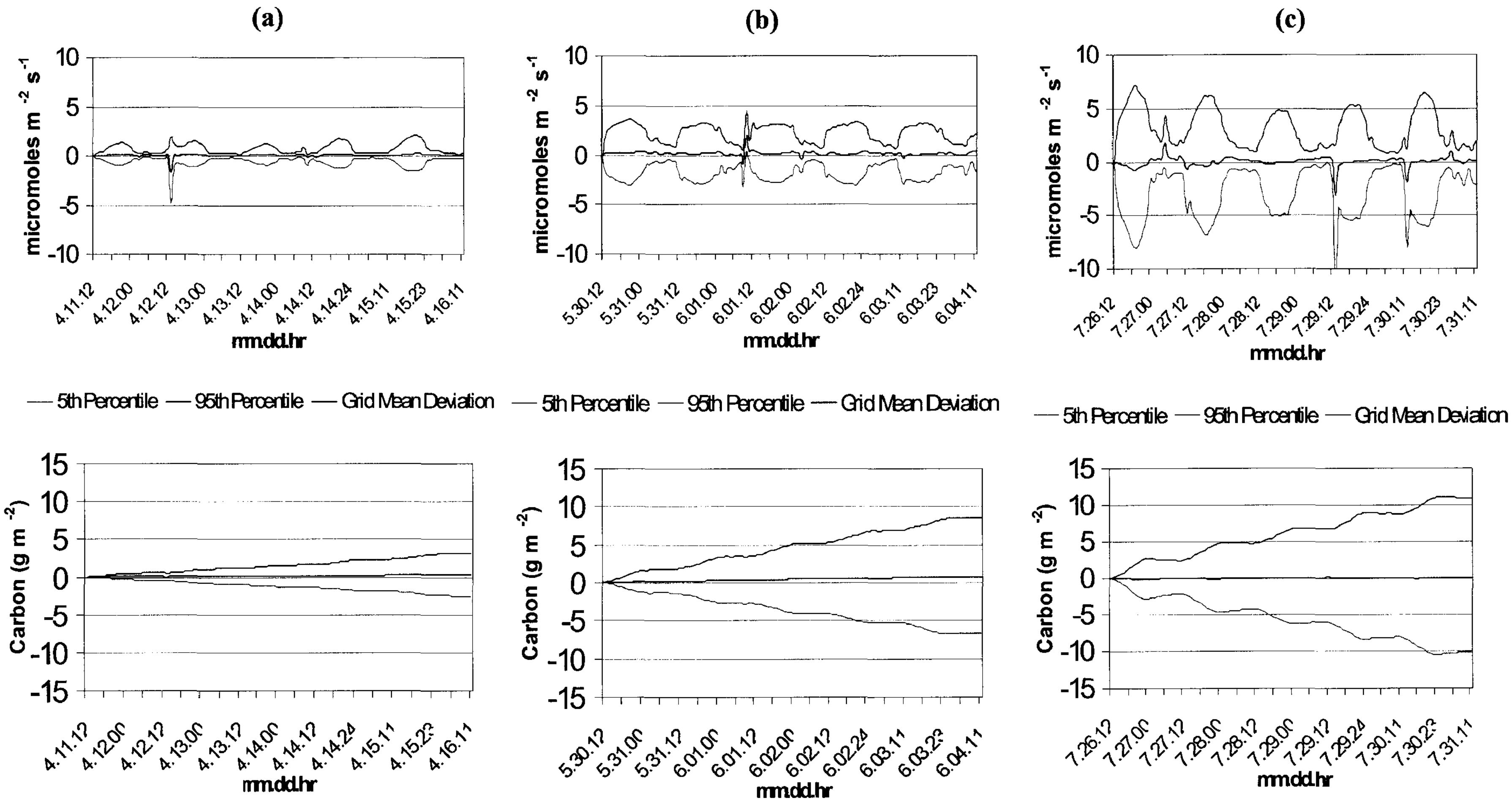


Figure 5.17 5th and 95th percentile deviations from the mean difference (top row) and cumulative mean deviations (bottom row) of net ecosystem exchange for a) April 11 – 16, b) May 30 – June 4 and c) July 26 – 31.

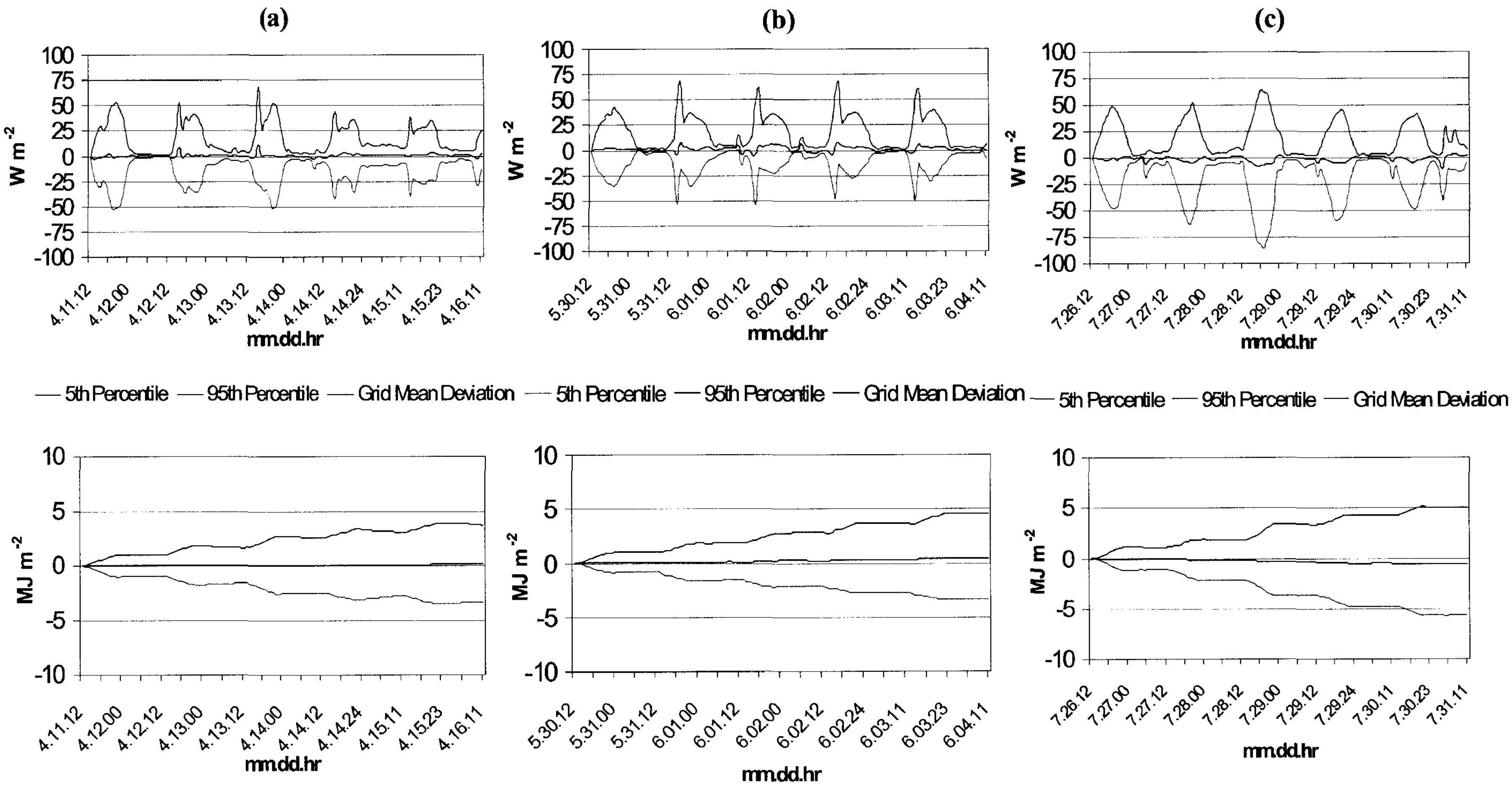


Figure 5.18 5th and 95th percentile deviations from the mean difference (top row) and cumulative mean deviations (bottom row) of sensible heat flux for a) April 11 – 16, b) May 30 – June 4 and c) July 26 – 31.

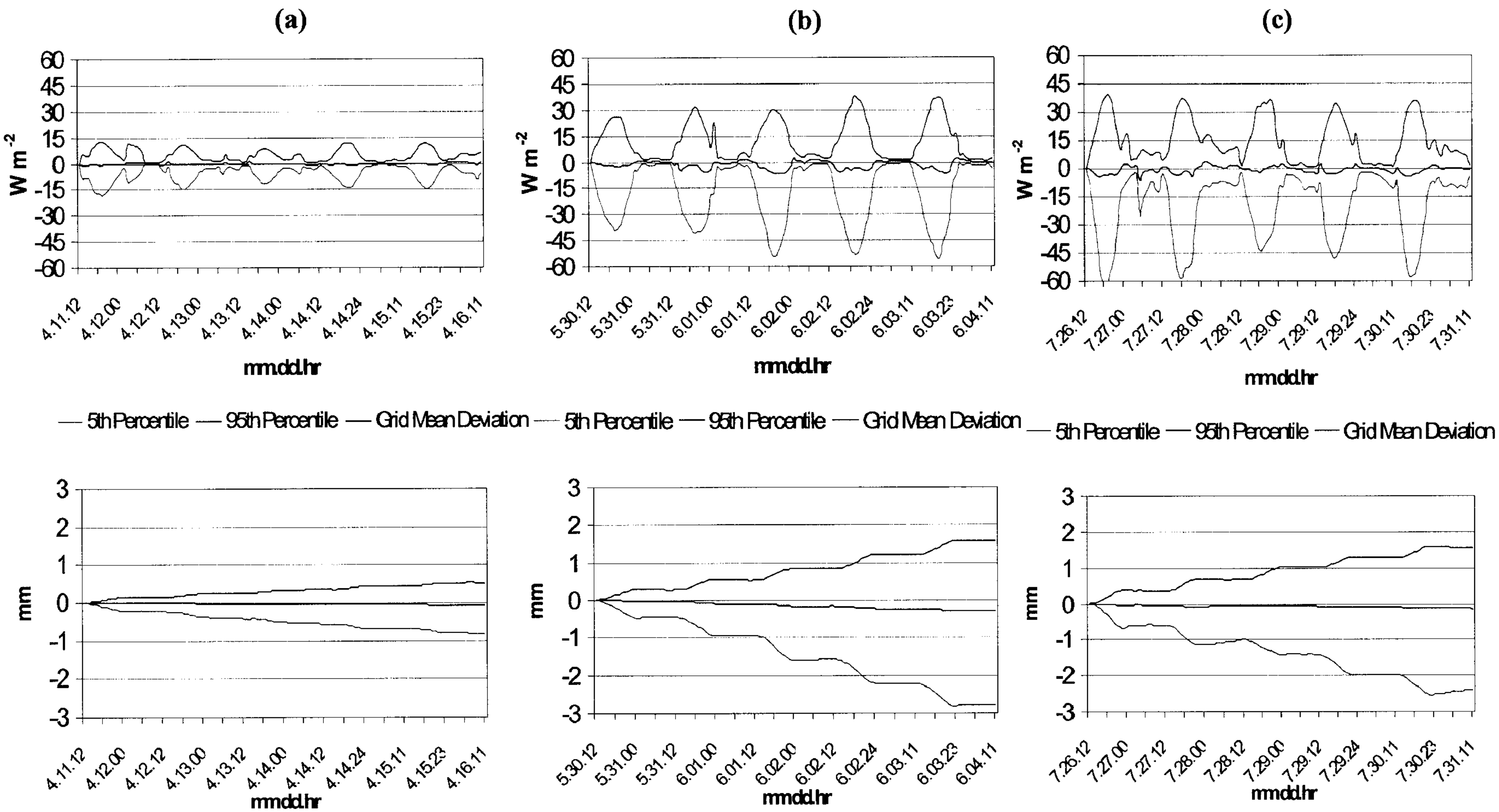


Figure 5.19 5th and 95th percentile deviations from the mean difference (top row) and cumulative mean deviations (bottom row) of water flux for a) April 11 – 16, b) May 30 – June 4 and c) July 26 – 31. Latent heat fluxes were converted to estimates of water using the latent heat of vaporization and the mean air temperature in each simulation period.

For sensible heat flux (Figure 5.18), the deviations and accumulated differences are closer in magnitude through time, though there are some obvious differences. For the late May/early June simulation there is a pronounced early morning difference in sensible heat flux between the two land surface representations. The April and July simulations have much less variability in terms of vegetation cover than the late May/early June simulations. In April, the deciduous leaves had not yet burst from their buds; in July, the vegetation was at full leaf area. In late May/early June the leaves were just bursting, resulting in a more heterogeneous surface as expressed through NDVI, prone to greater differential heating and cooling. The large deviations in Figure 5.18 may be caused by greater phenological heterogeneity, which is better captured at higher resolution.

For latent heat flux (Figure 5.19) the deviations and accumulated differences are quite variable. In April, when the land surface was quite cold, the deviations are also quite small and the accumulated differences amount to just less than a millimeter of water after 5 days. In late May/early June, as the land surface heats up, the differences are much larger resulting in almost more than two millimeters of water difference for the July simulations and slightly more for the late May/early June simulations. For all three time periods, the low heterogeneity simulations predict more water loss from the surface.

To put these differences in perspective, Figure 5.20 shows monthly the mean predictive uncertainty attributable to model parameterization calculated in Chapter 4 for the SiB2 model. For net ecosystem exchange, the errors due to spatial uncertainty (i.e. using a coarse resolution representation of the landscape instead of fine) in the late May/early June and July simulations are less than those due to parametric uncertainty. For example for the July simulations, the $5^{th}/95^{th}$ percentile error going from fine to coarse resolution was ~6 μmoles m^{-2} s^{-1} midday whereas the uncertainty due to parameterization was ~10 μmoles m^{-2} s^{-1}. For the April simulations, they are approximately the same magnitude (~2 μmoles m^{-2} s^{-1}). For sensible heat flux, in April and late May/early July the errors due to spatial uncertainty are smaller than those due to parametric uncertainty (e.g. 50 W m^{-2} vs. 79 W m^{-2} at midday for April), however for July the magnitudes of the differences due to spatial uncertainty approach those of the parametric uncertainty (e.g. 60 to 85 W m^{-2} vs. 82 W m^{-2} midday). For latent heat all the time periods showed the potential errors in latent heat flux due to using a coarse resolution representation of the landscape instead of fine were smaller than those due to parametric uncertainty (eg. for April midday 11 vs. 46 W m^{-2}; for July at midday 50 vs. 111 W m^{-2}. Late May/early June is harder to interpret since it straddles two months and so the potential errors due to spatial uncertainty might approach the uncertainty resulting from parameterization (eg. 50 W m^{-2} vs. 51 W m^{-2} midday in May and 98 W m^{-2} midday in June).

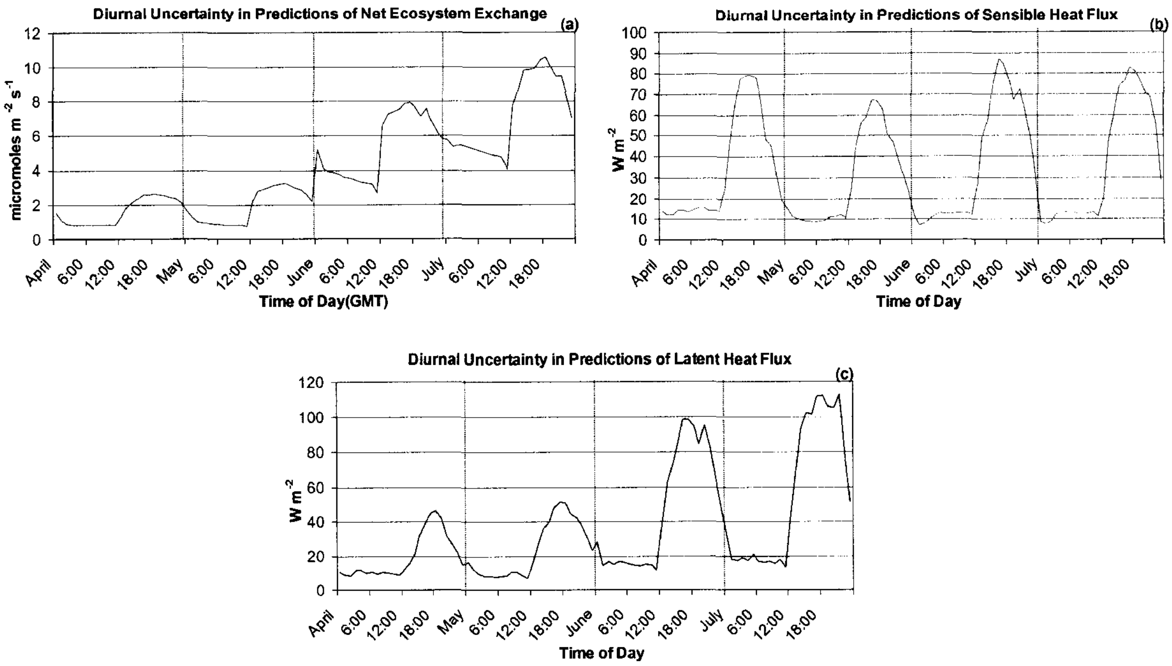


Figure 5.20 Diurnal predictive uncertainty constrained monthly on latent and sensible heat flux and net ecosystem exchange for April through July, 1997 (from Chapter 4). a) net ecosystem exchange, b) sensible heat flux and c) latent heat flux. The dotted lines represent the breaks between months.

5.5 Discussion

Heterogeneity of the land surface is scale dependent spatially as well as seasonally dependent. In this chapter, landscape heterogeneity was represented through a combination of satellite derived surface vegetation properties, spatially varying soil properties and a land cover biome classification. Even though the number of distinct vegetation classes was relatively small in the innermost grid of the simulations presented here (Table 5.3), a great deal of spatial variability in the landscape was nevertheless expressed through variability in the soil and the Normalized Difference Vegetation Index (NDVI) (Figures 5.11-5.13). Temporal variability in the simulations was exhibited through phenological changes expressed through NDVI and through the evolution of soil moisture and temperature dynamics through time. Loss of spatial heterogeneity was represented through aggregation of surface properties in space. This, in turn, reduced temporal variability in that the phenological differences expressed through NDVI between neighboring pixels at the fine scale were dampened. The scales of heterogeneity examined (1 km and 8 km) correspond to the spatial scales of widely available data sets, including two new global, long-term 8 km NDVI datasets (S. Los, C. Tucker personal communication).

The results presented here suggest that regional scale calculations of CO_2 flux are adequately captured with a moderate representation of landscape heterogeneity when compared to higher heterogeneity land surface representations. Similar results were found for both latent and sensible heat fluxes. However, there were small differences between the simulations that accumulated through time, suggesting that the differences are directional rather than random. For example, in the April and late May/early June simulations (excluding water), there was a 5% difference between the low and high heterogeneity simulations of net ecosystem exchange at the end of the 5-day simulation period. Though the fluxes at these times of year are small, the accumulation of differences (were they to continue) could lead to different predictions of source or sink status. It has been shown that growing season length is an important predictor of annual net ecosystem exchange (Goulden *et al.* 1996b). Small differences in surface representation during the shoulder seasons, the transition times between the beginning and end of the carbon uptake period, could push the source sink status of a nearly balanced ecosystem like WLEF in one direction or the other depending on the resolution of the land surface.

Even though on average the regional calculations of net ecosystem exchange, latent and sensible heat flux did not differ much between the representations of land surface heterogeneity, there were large differences in space and time at smaller scales (Figures 5.17-5.19). Thus, coarse scale representation of the land surface may lead to significant underestimation or overestimation of fluxes of carbon, water or heat at more local scales that could lead to changes in ecosystem water balance and carbon dynamics. If the finer-scale land surface representations are considered to be closer to the real world, then the impact of heterogeneity loss, particularly the water and heat balance components that interact strongly with the local atmosphere, is likely to become more significant over longer simulations as the trajectories diverge.

The deviations in the fluxes resulting from representing a fine resolution land surface coarsely are a measure of the spatial uncertainty in the land surface representation. Comparing these deviations to the uncertainty resulting from the parameterization of the SiB2 model calculated in Chapter 4, we can see that the magnitudes of the uncertainties in relation to each other vary by season and by flux. In general, the uncertainty due to representing a fine resolution land surface coarsely is less than the uncertainty due to the parameterization of the model itself, however in some instances (eg. July sensible heat flux, April net ecosystem exchange) they approach the same magnitudes.

It is likely that longer simulations would have continued the observed trends, as differential heating and cooling of the surface should cause greater deviations from the mean state. Unfortunately, the simulations presented in this chapter were limited in time by the need to choose clear sky conditions. Under these circumstances it is not reasonable to run coupled land surface–atmosphere simulations for long periods of time because it would result in an unreasonable radiation load reaching the surface causing accelerated drying of the surface and stress on the vegetation. A more natural course, and one seen in the observations, is that of repeated wetting and drying down periods and heterogeneous radiation due to clouds. An updated version of the coupled models will be available in the near future and extended simulations will be possible to see if these trends are indeed borne out.

In summary, the results reported in this chapter show that in general, uncertainty in surface flux estimates due to coarse scale representation of the land surface were less than the uncertainty attributable to the parameterization of the model as reported in Chapter 4. Nonetheless, local errors do appear to accumulate through time resulting in small differences between regional flux estimates and larger differences at individual points after only 5 days. This was particularly true for periods of rapid change such as the April and late May/early June simulations. Thus fine temporal and spatial resolution of land surface information is likely to be most critical in such periods, though how important will depend on whether the divergent trajectories are maintained over a longer periods of time. This is useful information for larger scale modelers, because it indicates that computing resources might be better spent resolving fine scale differences only in periods of rapidly changing surface conditions or in ecosystems where large land surface heterogeneity is maintained irrespective of seasonality, though that was not evaluated here. This result echoes that of Kaminski et al. (Kaminski *et al.* 2001) mentioned earlier, that resolving more basis regions was only useful when a certain amount of heterogeneity between the basis regions was maintained.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

The research presented in this dissertation addresses some of the key issues that carbon cycle science is currently facing and that will become more important within the framework of continental scaling efforts using model-data fusion such as the North American Carbon Program. In a series of analyses, sources of error, potential for bias and impacts of model and data choices in the context of regional carbon cycle modeling of temperate, forested ecosystems were evaluated. A comprehensive sensitivity analyses was conducted to evaluate the sensitivity of a commonly used, complex biophysical land surface model to its parameterization through time. Next, the uncertainty of land surface fluxes predicted by the model was calculated. Finally the potential for uncertainty due to the spatial representation of the land surface was evaluated. The results of the analyses presented in the previous chapters have important implications for carbon cycle modeling in general and carbon cycle data assimilation in particular.

In conducting the sensitivity analysis, in which tens of thousands of simulations were run with randomly varying but realistic parameter sets, it was found that there was an irreducible level of mismatch between the simulated and observed fluxes. It was not possible to predict fluxes of latent and sensible heat and net ecosystem exchange beyond certain limits, which varied by month and by flux. This minimum possible error reflects a difference between simulated and observed fluxes that cannot be overcome through optimization due to variability (or noise) in the observations and/or structural problems with the model. It is a measure of uncertainty that reflects the expected degree of hourly model-data mismatch over periods of months to a year. In the context of regional inversions, the minimum error can be used to estimate prior model uncertainty where modeled fluxes are used or an estimate of uncertainty for assimilation of measured fluxes.

The sensitivity analysis also revealed that of the 46 parameters evaluated, on average fewer than half were found to be influential to model results. This is useful information in that clearly non-influential parameters are ones that should not be optimized in a data analysis system. However, the number and identity of influential parameters varied considerably by month. In most cases the dominant parameters were a reflection of the dominant processes during that time period. For example, parameters related to photosynthesis were more influential during the growing season and parameters related to energy balance were more influential at other times of the year. The greatest sensitivity to parameters, both in the magnitude of the sensitivity as well as in the number of parameters, occurred in the shoulder seasons. The shoulder seasons are those periods of rapid land surface change surrounding leaf-on in the spring and leaf-off in the autumn. During these periods a combination of parameters reflecting energy balance, the temperature response of photosynthesis and the aerodynamic resistance of the land surface were most influential. There were also patterns of interest within seasons. For example parameters such as the rubisco velocity of a leaf in full sun and the slope of the Ball-Berry photosynthesis-conductance relationship, which are usually assumed to be time-invariant in models such as SiB2, underwent subtle shifts during the growing season. It is possible this reflects the changing photosynthetic capacity of leaves throughout the growing season. Thus, the variability of model sensitivity through time suggests that any optimization on tower flux data should account for and allow for variability at sub annual time scales in order to capture the most information.

A surprising result was the lack of sensitivity to the parameters controlling heterotrophic respiration. None of the parameters varied in this study that influence the rate of heterotrophic respiration in SiB2 were particularly influential. Though the temperature response function (Q_{10}) was not varied, it was also shown that it likely would not have been influential either. In part, this is because the variability in the observations themselves overwhelmed the relatively more subtle response to these parameters.

The results of the sensitivity analysis also emphasized the risk involved in undirected model optimization. Parameter interactions in complex models such as SiB2 are very non-linear and thus complicated to interpret. Compensation among parameters led to many parameter sets that did an equally good job of predicting fluxes, particularly when the model was being tested against multiple fluxes. This is a serious consideration for optimization studies, suggesting that uncertainty resulting from the parameterization of the model should be incorporated in any model-data fusion exercise. Besides compensating for one another, the parameters also compensate for imperfect driver data and model structural problems. In some instances, the parameter distributions resulting from the sensitivity analysis were not biologically realistic with respect to observed conditions in the real world, though they produced the best fluxes. Thus, too highly automated optimization procedures should be viewed with caution, especially since it can be difficult to determine whether better parameter values or better model structure are needed.

The uncertainty analysis showed that overall, model structure and logic was quite good for net ecosystem exchange. More than 50% of the observations consistently fell within the predictive uncertainty bounds. For latent and sensible heat flux, however, far fewer observations consistently fell with in the uncertainty bounds. It is possible that heat fluxes are simply more variable across the landscape than net ecosystem exchange or that the mean field parameterization and functional group representation adequate for net ecosystem exchange may be less so for the heat fluxes.

When comparing observational constraints on the model at monthly versus yearly time scales, it was found that latent and sensible heat flux and daytime net ecosystem exchange were better constrained at the shorter, monthly time period. Nighttime net ecosystem exchange was better constrained annually. These results indicate, again, that the time scale of the observed process should be taken into consideration when considered within a data assimilation framework. Even though fluxes of latent and sensible heat and net ecosystem exchange are measured simultaneously, the processes they represent have different temporal signatures. Further, these results also suggest that rather than overwhelm the problem of optimization with observations, there are ways to maximize the value of the data, such as breaking it down into shorter time periods for some processes and not for others.

The predictive uncertainty bounds also indicated systematic biases in predicted fluxes. Systematic biases can be caused by incorrect process representation in the model or systematic errors in the observations, such as the much-noted difficulty in observing nighttime net ecosystem exchange. The results presented here showed that uncertainties were quite large relative to fluxes during the shoulder seasons. This was attributed to the estimation of time-varying vegetation parameters with a single, monthly maximum value composited satellite product and linear interpolation between months. When model parameterization was allowed to vary monthly, much of the uncertainty was reduced. However, there were still within-month biases indicating that information at a finer time scale resolution could reduce uncertainties further for these sensitive periods. The model also exhibited very little variability in predicted nighttime fluxes of net ecosystem exchange compared to the observations, such that nighttime observations were far less likely to fall within the uncertainty bounds than daytime fluxes.

Regional scale calculations of CO_2 flux were adequately captured with a moderate representation (8 km) of landscape heterogeneity when compared to higher heterogeneity (1 km) land surface representations. Similar results were found for both latent and sensible heat fluxes. However, there were small differences between the simulations that accumulated through time, suggesting that the differences are systematic rather than random. The accumulation of differences (were they to continue) could lead to different predictions of source or sink status, particularly in a nearly balanced ecosystem like WLEF. For net ecosystem exchange, these differences were more pronounced in the shoulder season simulations.

At more local scales, there were large differences between predicted fluxes in space and time. If the finer-scale land surface representations are considered to be closer to the real world, then the local impact of heterogeneity loss, particularly the water and heat balance components that interact strongly with the local atmosphere, is likely to become more significant over longer simulations as the trajectories diverge. The deviations in the fluxes resulting from representing a fine resolution land surface coarsely were compared to the uncertainty resulting from the parameterization of the SiB2 model calculated in Chapter 4. In general, the uncertainty due to representing a fine resolution land surface coarsely was less than the uncertainty due to the parameterization of the model itself. However in some instances they approached the same magnitudes. For regional atmospheric inversions, these results suggest that fine temporal and spatial resolution of land surface information is likely to be most critical during time periods of rapid change. Further, greater reductions in uncertainty for the model predictions presented here might be better recovered by improving parameter uncertainty and/or model structure, than by improving the representation of land surface heterogeneity, at least initially.

Many of the results presented in this dissertation may be used to develop recommendations regarding which directions might be most profitably pursued to reduce uncertainty in regional carbon cycle modeling. Improvement in the space-time flux patterns used by atmospheric inversion models will reduce the bias and uncertainty in the *a posteriori* fluxes, increasing the robustness of the flux estimates. Improving these space-time flux patterns through carbon cycle data assimilation ensures that the improvement is not only in the resultant large-scale flux estimates, but also in the predictive and explanatory power of the terrestrial biospheric models themselves. Model data fusion is non-trivial, requiring the synergistic development of multiple models along with improved observations and significant amounts of computing power. Sensitivity and uncertainty analyses such as the ones presented here help to define the areas where the most benefit can be gained.

The ideas explored in this dissertation are just the beginning of what will be needed in the future. The coupled model simulations would benefit from both longer runs and the inclusion of cloud microphysics so that the mechanisms behind any differences caused by the degree of expressed land surface heterogeneity could be more fully explored. It would also be very useful to conduct a similar sensitivity and uncertainty analysis with analyzed weather fields to explore the effects of the driver data. One problem that modelers often run into is the lack of appropriate data with which to run their models. Even though there are quite a large number of flux tower sites in North America, not all of them have the requisite filled driver data. To calculate more broadly applicable estimates of uncertainty using the techniques presented here, more sites would need to be evaluated. Thus characterizing the uncertainty related to the driver data would be a useful way to expand the number of possible sites.

REFERENCES

Alapaty, K., S. Raman and D. S. Niyogi (1997). Uncertainty in the Specification of Surface Characteristics: a Study of Prediction Errors in the Boundary Layer. *Boundary Layer Meteorology* **82**: 473-500.

Avissar, R. (1995). Scaling land-atmosphere interactions: an atmospheric modeling perspective. *Hydrological Processes* **9**: 679-695.

Avissar, R., P. L. S. Dias, M. Dias and C. Nobre (2002). The Large-Scale Biosphere-atmosphere Experiment in Amazonia (LBA): Insights and future research needs. *Journal of Geophysical Research-Atmospheres* **107**(D20).

Avissar, R. and R. A. Pielke (1991). Impact of stomatal control on mesoscale atmospheric circulations. *Agricultural and forest meteorology*(54): 353-372.

Baker, I., A. S. Denning, N. P. Hanan, L. Prihodko, M. Uliasz, P.-L. Vidale, K. J. Davis and P. S. Bakwin (2003). Simulated and observed fluxes of sensible and latent heat and CO2 at the WLEF-TV tower using SiB2.5. *Global Change Biology* **9**: 1262-1277.

Bakwin, P. S., P. P. Tans, D. F. Hurst and C. L. Zhao (1998). Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program. *Tellus Series B-Chemical and Physical Meteorology* **50**(5): 401-415.

Baldocchi, D., E. Falge, L. H. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R. Evans, J. Fuentes, A. Goldstein, G. Katul, B. Law, X. H. Lee, Y. Malhi, T. Meyers, W. Munger, W. Oechel, K. T. P. U, K. Pilegaard, H. P. Schmid, R. Valentini, S. Verma, T. Vesala, K. Wilson and S. Wofsy (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society* **82**(11): 2415-2434.

Baldocchi, D. D. (2003). Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology* **9**(4): 479-492.

Ball, J. T., I. E. Woodrow and J. A. Berry (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. Progress in photosynthesis research. J. Biggins. Dordrecht, Martinus Nijhoff. **4:** 221-224.

Bastidas, L. A., H. V. Gupta, S. Sorooshian, W. J. Shuttleworth and Z. L. Yang (1999). Sensitivity analysis of a land surface scheme using multicriteria methods. *Journal of Geophysical Research* **104**(D16): 19,481-19,490.

Berger, B. W., K. J. Davis, C. X. Yi, P. S. Bakwin and C. L. Zhao (2001). Long-term carbon dioxide fluxes from a very tall tower in a northern forest: Flux measurement methodology. *Journal of Atmospheric and Oceanic Technology* **18**(4): 529-542.

Beven, K. J. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Research* **16**: 41-51.

Beven, K. J. and A. Binley (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes* **6**: 279-298.

Beven, K. J. and J. Freer (2001). Equifinality, data assimilation and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* **249**: 11-29.

Bonan, G. B. (1996). A Land Surface Model (LSM version 1.0) for Ecological, Hydrological and Atmospheric Studies: Technical Description and User's Guids. Boulder, National Center for Atmospheric Research: 148.

Bounoua, L., R. DeFries, G. J. Collatz, P. J. Sellers and H. Khan (2002). Effects of land cover conversion on surface climate. *Climatic Change* **52**(1-2): 29-64.

Bousquet, P., P. Peylin, P. Ciais, C. Le Quere, P. Friedlingstein and P. Tans (2000). Regional changes in carbon dioxide fluxes of land and oceans since 1980. *Science* **190**: 1342-1346.

Brazier, R. E., K. J. Beven, J. Freer and J. S. Rowan (2000). Equifinality and uncertainty in physically based soil erosion models: application of the GLUE methodology to WEPP-the water erosion prediction project-for sites in the UK and USA. *Earth Surface Processes and Landforms* **25**: 825-845.

Brazier, R. E., K. J. Beven, S. G. Nathony and J. S. Rowan (2001). Implications of model uncertainty for the mapping of hillslope-scale soil erosion predictions. *Earth Surface Processes and Landforms* **26**: 1333-1352.

Brewer, C. A. and L. J. Gross (2003). Training ecologists to think with uncertainty in mind. *Ecology* **86**(6): 1412-1414.

Burrows, S. N., S. T. Gower, M. K. Clayton, D. S. Mackay, D. E. Ahl, J. M. Norman and G. Diak (2002). Application of geostatistics to characterize LAI for flux towers to landscapes. *Ecosystems* **5**: 667-679.

Calder, C., M. Lavine, P. Muller and J. S. Clark (2003). Incorporating multiple sources of stochasticity into dynamic population models. *Ecology* **84**(6): 1395-1402.

Charney, J. G. (1975). Dynamics of deserts and drought in the Sahel. *Quarterly Journal of the Royal Meteorological Society* **101**: 193-202.

Chase, T. N., R. A. Pielke Sr., T. G. Kittel, J. S. Baron and T. J. Stohlgren (1999). Potential impacts on Colorado Rocky Mountain weather due to land use changes on the adjacent great plains. *Journal of Geophysical Research* **104**(D14): 16673-16690.

Chen, F., D. Yates, H. Nagai, M. A. LeMone, K. Ikeda and R. L. Grossman (2003). Land surface heterogeneity in the Cooperative Atmosphere Surface Exchange Study (CASES-97). Part I: Comparing modeled surface flux maps with surface-flux tower and aircraft measurements. *Journal of Hydrometeorology* **4**: 196-218.

Ciais, P., P. P. Tans, M. Trolier, J. W. C. White and R. J. Francey (1995). A Large northern hemisphere terrestrial carbon sink indicated by the $^{13}C/^{12}C$ ratio of Atmospheric CO_2. *Science* **269**: 1098-1102.

Clapp, R. B. and G. M. Hornberger (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research* **14**(4): 601-604.

Clark, J. S. (2003). Uncertainty and variability in demography and population growth: a hierarchical approach. *Ecology* **84**(6): 1370-1381.

Clark, J. S., J. Mohan, M. Dietze and I. Ibanez (2003). Coexistence: How to identify trophic tradeoffs. *Ecology* **84**(1): 17-31.

Colello, G. D., C. Grivet, P. J. Sellers and J. A. Berry (1998). Modeling of energy, water, and CO2 flux in a temperate grassland ecosystem with SiB2: May-October 1987. *Journal of the Atmospheric Sciences* **55**(7): 1141-1169.

Collatz, G. J., J. T. Ball, C. Grivet and J. A. Berry (1991). Physiological and Environmental-Regulation of Stomatal Conductance, Photosynthesis and Transpiration - a Model That Includes a Laminar Boundary-Layer. *Agricultural and Forest Meteorology* **54**(2-4): 107-136.

Collatz, G. J., L. Bounoua, S. O. Los, D. A. Randall, I. Y. Fung and P. J. Sellers (2000). A mechanism for the influence of vegetation on the response of the diurnal temperature range to changing climate. *Geophysical Research Letters* **27**(20): 3381-3384.

Collatz, G. J., M. Ribascarbo and J. A. Berry (1992). Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants. *Australian Journal of Plant Physiology* **19**(5): 519-538.

Collins, D. C. and R. A. Avissar (1994). An evaluation with the Fourier Amplitude Sensitivity Test (FAST) of which land surface parameters are of greatest importance in atmospheric modelling. *Journal of Climate* **7**: 681-703.

Cosby, B. J., G. M. Hornberger, R. B. Clapp and T. R. Ginn (1984). A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research* **20**(6): 682-690.

Cotton, W. R., R. A. Pielke, R. L. Walko, G. E. Liston, C. J. Tremback, H. Jiang, R. L. McAnelly, J. Y. Harrington, M. E. Nicholls, G. G. Carrio and J. P. McFadden (2003). RAMS 2001: Current status and future directions. *Meteorology and Atmospheric Physics* **82**(1-4): 5-29.

Cotton, W. R. and R. A. Pielke Sr. (1995). Human Impacts on Weather and Climate. Cambridge, Cambridge Uiversity Press.

Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall and I. J. Toterdell (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**: 184-187.

Dale, V. H., H. I. Jager, R. H. Gardner and A. E. Rosen (1988). Using sensitivity and uncertainty analyses to improve predictions of broad-scale forest development. *Ecological Modelling* **42**: 165-178.

Davis, K. J., P. S. Bakwin, C. Yi, B. W. Berger, C. Zhao, R. M. Teclaw and J. G. Isebrands (2003). The annual cycles of CO2 and H2O exchange over a northern mixed forest as observed form a very tall tower. *Global Change Biology* **9**: 1278-1293.

Denning, A. S., G. J. Collatz, C. G. Zhang, D. A. Randall, J. A. Berry, P. J. Sellers, G. D. Colello and D. A. Dazlich (1996a). Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model .1. Surface carbon fluxes. *Tellus Series B-Chemical and Physical Meteorology* **48**(4): 521-542.

Denning, A. S., I. Y. Fung and D. Randall (1995). Latitudinal gradient of atmospheric CO_2 due to seasonal exchange with land biota. *Nature* **376**: 240-243.

Denning, A. S., M. Nicholls, L. Prihodko, I. Baker, P.-L. Vidale, K. J. Davis and P. S. Bakwin (2003). Simulated variations in atmospheric CO2 over a Wisconsin forest using a coupled ecosystem-atmosphere model. *Global Change Biology* **9**: 1241-1250.

Denning, A. S., R. Oren, D. McGuire, C. Sabine, S. C. Doney, K. Paustian, M. Torn, L. Dilling, L. Heath, P. Tans, S. Wofsy, R. Cook, A. Andrews, G. P. Asner, J. Baker, P. Bakwin, R. A. Birdsey, D. Crisp, K. Davis, C. Field, C. Gerbig, H. Mayeus, C. McClain, B. McKee, C. Miller, S. Pawson, J. Randerson, J. Reilly, S. Running, S. Saleska, R. Stallard, E. T. Sundquist, S. Ustin and S. Verma (2004). Science implementation strategy for the North American Carbon Program, Prepared for the Carbon Cycle Science Steering Group and the Interagency Working Group on Carbon: 75.

Denning, A. S., D. A. Randall, G. J. Collatz and P. J. Sellers (1996b). Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model .2. Simulated CO2 concentrations. *Tellus Series B-Chemical and Physical Meteorology* **48**(4): 543-567.

Dennis, B. (1996). Discussion: Should ecologists become Bayesians. *Ecological Applications* **6**(4): 1095-1103.

Desjardins, R. L., J. I. MacPherson, H. H. Neumann, G. Den Hartog and P. M. Schuepp (1995). Flux estimates of latent and sensible heat, carbon dioxide and ozone using an aircraft-tower combination. *Atmospheric Environment* **29**(21): 3147-3158.

Eastman, J. L., R. Pielke Sr. and D. J. McDonald (1998). Calibration of soil moisture for large eddy simulations over the FIFE area. *Journal of the Atmospheric Sciences* **55**: 1131-1140.

Eidenshink, J. C. and J. L. Faundeen (1994). The 1-Km AVHRR global land data set: first stages in implementation. *International Journal of Remote Sensing* **15**: 3443-3462.

El Maayar, M., D. T. Price, T. A. Black, E. R. Humphreys and E.-M. Jork (2002). Sensitivity tests of the integrated biosphere simulator to soil and vegetation charachteristics in a pacific coastal coniferous forest. *Atmosphere-Ocean* **40**(3): 313-332.

Ellison, A. M. (1996). An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications* **6**(4): 1036-1046.

Engelen, R. J., A. S. Denning and K. R. Gurney (2002). On error estimation in atmospheric CO2 inversions. *Journal of Geophysical Research-Atmospheres* **107**(D22).

Enting, I. G., C. M. Trudinger and R. J. Francey (1995). A synthesis inversion of the concentration and d13C of atmospheric CO2. *Tellus* **47B**: 35-52.

Errico, R. M. and T. Vukicevic (1992). Sensitivity analysis using and adjoint of the PSU-NCAR mesoscale model. *Monthly Weather Review* **120**: 1644-1660.

Ewers, B. E., D. S. Mackay, S. T. Gower, D. E. Ahl, S. N. Burrows and S. S. Samanta (2002). Tree species effects on stand transpiration in northern Wisconsin. *Water Resources Research* **38**(7).

Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi and P. Tans (1998). A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. *Science* **282**: 442-446.

Farquhar, G. D., S. V. Caemmerer and J. A. Berry (1980). A Biochemical-Model of Photosynthetic Co2 Assimilation in Leaves of C-3 Species. *Planta* **149**(1): 78-90.

Fedra, K., G. Van Straten and M. B. Beck (1981). Uncertainty and arbitrariness in ecosystems modelling: a lake modelling example. *Ecological Modelling* **13**: 87-110.

Field, C. and H. A. Mooney (1986). The photosynthesis-nitrogen relationship in wild plants. On the Economy of Plant Form and Function. T. J. Givnish. Cambridge, Cambridge University Press: 25-55.

Franks, S. W. (1998). An evaluation of single and multiple objective SVAT model conditioning schemes: parametric, predictive and extrapolative uncertainty. Callagghan, New South Wales, The University of Newcastle, Department of Civil, Surveying and Environmental Engineering: 34.

Franks, S. W. and K. J. Beven (1997). Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. *Journal of Geophysical Research* **102**(D20): 23,991-23,999.

Franks, S. W. and K. J. Beven (1999). Conditioning a multiple-patch SVAT model using uncertain time-space estimates of latent heat fluxes as inferred from remotely sensed data. *Water Resources Research* **35**(9): 2751-2761.

Franks, S. W., K. J. Beven and J. H. C. Gash (1999). Multi-objective conditioning of a simple SVAT model. *Hydrology and Earth System Sciences* **3**(4): 477-489.

Franks, S. W., K. J. Beven, P. F. Quinn and I. R. Wright (1997). On the sensitivity of soil-vegetation-atmosphere transfer (SVAT) schemes: equifinality and the problem of robust calibration. *Agricultural and Forest Meteorology* **86**: 63-75.

Franks, S. W., P. Gineste, K. J. Beven and P. Merot (1998). On constraining the predictions of a distributed mode: the incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resources Research* **34**(4): 787-797.

Freer, J. and K. J. Beven (1996). Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resources Research* **32**(4): 2161-2173.

Friedl, M. A. (1996). Relationships among remotely sensed data, surface energy balance, and area-averaged fluxes over partially vegetated land surfaces. *Journal of Applied Meteorology* **35**: 2091-2103.

Gardner, R. H., R. V. O'Neill, J.B. Mankin, J.H. Carney (1981). A comparison of sensitivity analysis and error analysis based on a stream ecosystem model. *Ecological Modelling* **12**: 173-190.

Geels, C., S. C. Doney, R. Dargaville, J. Brandt and J. H. Christensen (2004). Investigating the sources and sinks of synoptic variability in atmospheric CO2 measurements over the Northern Hemisphere continents: a regional model study. *Tellus* **56B**: 35-40.

Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin and C. A. Grainger (2003a). Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms. *Journal of Geophysical Research-Atmospheres* **108**(D24).

Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin and C. A. Grainger (2003b). Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *Journal of Geophysical Research-Atmospheres* **108**(D24).

Goulden, M., J. W. Munger, S. Fan, B. C. Daube and S. Wofsy (1996a). Measurements of carbon sequestration by long term eddy covariance: methods and a critical evaluation of accuracy. *Global Change Biology* **2**: 169-182.

Goulden, M. J., W. J. Munger, S. M. Fan, B. C. Daube and S. C. Wofsy (1996b). Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability. *Science* **271**: 1576-1578.

Gupta, H. V., L. Bastidas, S. Sorooshian, W. J. Shuttleworth and Z.-L. Yang (1999). Parameter estimation of a land surface scheme using multicriteria methods. *Journal of Geophysical Research* **104**(D16): 19491-19503.

Gupta, H. V., S. Sorooshian and P. O. Yapo (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research* **34**(4): 751-763.

Gurney, K. R., R. M. Law, A. S. Denning, P. J. Rayner, D. Baker, P. Bousquet, L. Bruhwiler, Y. H. Chen, P. Ciais, S. Fan, I. Y. Fung, M. Gloor, M. Heimann, K. Higuchi, J. John, T. Maki, S. Maksyutov, K. Masarie, P. Peylin, M. Prather, B. C. Pak, J. Randerson, J. Sarmiento, S. Taguchi, T. Takahashi and C. W. Yuen (2002). Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models. *Nature* **415**(6872): 626-630.

Hall, F. G., K. F. Huemmrich, S. J. Goetz, P. J. Sellers and J. E. Nickeson (1992). Satellite remote sensing of surface energy balance: Successes, failures and unresolved issues in FIFE. *Journal of Geophysical Research* **97**: 19061-19089.

Hall, F. G. and P. J. Sellers (1995). First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE) in 1995. *Journal of Geophysical Research-Atmospheres* **100**(D12): 25383-25395.

Hall, M. C. G., D. G. Cacuci and M. E. Schlesinger (1982). Sensitivity analysis of a radiative-convective model by the adjoint method. *Journal of the Atmospheric Sciences* **39**: 2038-2050.

Hallgren, W. S. and A. J. Pitman (2000). The Uncertainty in Simulations by a Global Biome Model (BIOME3) to Alternative Parameter Values. *Global Change Biology* **6**: 483-495.

Hanan, N. P., J. A. Berry, S. Verma, E. A. Walter-Shea, A. E. Suyker, G. G. Burba and A. S. Denning (in prep). Model analyses of biosphere-atmosphere exchange of CO2, water and energy in Great Plains tallgrass prairie and wheat ecosystems.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend and R. Sohlberg (2000). Global land cover classification at 1 Km spatial resolution using a classification tree approach. *Journal of Remote Sensing* **21**(6): 1331-1364.

Harrington, J. Y. (1997). The effects of radiative and microphysical processes on simulated warm and transition season Arctic stratus. Department of Atmospheric Science. Fort Collins, Colorado State University: 289.

Henderson-Sellers, A. (1993). A factorial assessment of the sensitivity of the BATS land-surface parameterization scheme. *Journal of Climate* **6**: 227-247.

Hornberger, G. M. and R. C. Spear (1981). An approach to the preliminary analysis of environmental systems. *Journal of Environmental Management* **12**: 7-18.

Houghton, J. T., Y. Ding, M. Griggs, M. Noguer, P. J. van der Linden and D. Xiaosu, Eds. (2001). Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, Cambridge University Press.

Houghton, J. T., G. J. Jenkins and J. J. Ephraums, Eds. (1990). Climate Change: The IPCC Scientific Assessment. Cambridge, Cambridge University Press.

Houghton, R. A. (2003). Why are estimates of the terrestrial carbon balance so different? *Global Change Biology* **9**: 500-509.

Hu, Z. and S. Islam (1997). Effects of spatial variability on the scaling of land surface parameterizations. *Boundary Layer Meteorology* **83**(3): 441-461.

Idso, S. B. (1981). A set of equations for full spectrum and 8mm to 14mm and 10.5mm to 12.5mm thermal radiation for cloudless skies. *Water Resources Research* **17**: 295-304.

James, G. and R. C. James (1976). Mathematics Dictionary. New York, Van Nostrand Reinhold Company.

Jones, H. G. (1992). Plants and microclimate: a quantitative approach to environmental plant physiology. Cambridge, Cambridge University Press.

Kalnay, E. (2003). Atmospheric Modeling, Data Assimilation and Predictability. Cambridge, Cambridge University Press.

Kaminski, T., R. Giering, M. Scholze, P. Rayner and W. Knorr (2003). An example of an automatic differentiation-based modelling system. Computational Science and Its Applications - Iccsa 2003, Pt 2, Proceedings. **2668:** 95-104.

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

Kaminski, T., P. J. Rayner, M. Heimann and I. G. Enting (2001). On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research-Atmospheres* **106**(D5): 4703-4715.

Kheshgi, H. S., A. K. Jain and D. J. Wuebbles (1999). Model-based estimation of the global carbon budget and its uncertainty from carbon dioxide and carbon isotope records. *Journal of Geophysical Research* **104**(D24): 31,127-31,143.

Kimball, J. S., S. Running and S. S. Saatchi (1999). Sensitivity of boreal forest regional water flux and net primary production simulations to sub-grid-scale land cover complexity. *Journal of Geophysical Research* **104**(D22): 27,789-27,801.

Kjellstrom, E., K. Holmen, K. Eneroth and M. Engardt (2002). Summertime Siberian CO2 simulations with regional transport model MATCH: a feasibility study of carbon uptake calculations from EUROSIB data. *Tellus* **54B**: 834-849.

Klemp, J. B. and R. B. Wilhelmson (1978). Simulation of 3-Dimensional Convective Storm Dynamics. *Journal of the Atmospheric Sciences* **35**(6): 1070-1096.

Kremer, J. N. (1983). Ecological implications of parameter uncertainty in stochastic simulation. *Ecological Modelling* **18**: 187-207.

Kustas, W. P. and T. J. Jackson (1999). The impact on area-averaged heat fluxes from using remotely sensed data at different resolutions: a case study with Washita '92 data. *Water Resources Research* **35**(5): 1539-1550.

Lavigne, M. B., M. G. Ryan, D. E. Anderson, D. D. Baldocchi, P. M. Crill, D. R. Fitzjarrald, M. L. Goulden, S. T. Gower, J. M. Massheder, J. H. McCaughey, M. Rayment and R. G. Striegl (1997). Comparing nocturnal eddy covariance measurements to estimates of ecosystem respiration made by scaling chamber measurements at six coniferous boreal sites. *Journal of Geophysical Research-Atmospheres* **102**(D24): 28977-28985.

Law, B. E., D. Baldocchi and P. M. Anthioni (1999). Below canopy and soil CO2 fluxes in a ponderosa pine forest. *Agricultural and Forest Meteorology* **94**: 171-188.

Law, B. E., E. Falge, L. Gu, D. D. Baldocchi, P. Bakwin, P. Berbigier, K. Davis, A. J. Dolman, M. Falk, J. D. Fuentes, A. Goldstein, A. Granier, A. Grelle, D. Hollinger, I. A. Janssens, P. Jarvis, N. O. Jensen, G. Katul, Y. Mahli, G. Matteucci, T. Meyers, R. Monson, W. Munger, W. Oechel, R. Olson, K. Pilegaard, K. T. Paw, H. Thorgeirsson, R. Valentini, S. Verma, T. Vesala, K. Wilson and S. Wofsy (2002). Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. *Agricultural and Forest Meteorology* **113**(1-4): 97-120.

Law, R. M., Y. H. Chen and K. R. Gurney (2003a). TransCom 3 CO2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices. *Tellus Series B-Chemical and Physical Meteorology* **55**(2): 580-595.

Law, R. M., P. J. Rayner, L. P. Steele and I. G. Enting (2003b). Data and modelling requirements for CO2 inversions using high-frequency data. *Tellus Series B-Chemical and Physical Meteorology* **55**(2): 512-521.

Los, S. O., G. J. Collatz, P. J. Sellers, C. M. Malmstrom, N. H. Pollack, R. S. DeFries, L. Bounoua, M. T. Parris, C. J. Tucker and D. A. Dazlich (2000). A global 9-yr biophysical land surface dataset from NOAA AVHRR data. *Journal of Hydrometeorology* **1**(2): 183-199.

Lu, L. and W. J. Shuttleworth (2002). Incorporating NDVI-Derived LAI into the climate version of RAMS and it's impact on regional climate. *Journal of Hydrometeorology* **3**: 347-362.

Mackay, D. S., D. E. Ahl, B. E. Ewers, S. T. Gower, S. N. Burrows, S. Samanta and K. J. Davis (2002). Effects of aggregated classification of forest composition on estimates of evapotranspiration in a northern Wisconsin forest. *Global Change Biology* **8**(12): 1253-1266.

Mahrt, L. and D. Vickers (2002). Relationship of area-averaged carbon dioxide and water vapor fluxes to atmospheric variables. *Agricultural and Forest Meteorology* **112**: 195-202.

Margulis, S. A. and D. Entekhabi (2001a). A coupled land surface-boundary layer model and its adjoint. *Journal of Hydrometeorology* **2**: 274-296.

Margulis, S. A. and D. Entekhabi (2001b). Feedback between the land surface energy balance and atmospheric boundary layer diagnosed through a model and its adjoint. *Journal of Hydrometeorology* **2**: 599-620.

Martin, L. (1965). The Physical Geography of Wisconsin. Madison and Milwaukee, The University of Wisconsin Press.

McCarthy, J. J., O. F. Canziani, N. A. Leary, D. J. Dokken and K. S. White, Eds. (2001). Climate Change 2001: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, Cambridge University Press.

Meixner, T., H. V. Gupta, L. Bastidas and R. Bales (1999). Sensitivity analysis using mass flux and concentration. *Hydrological Processes* **13**(2233-2244).

Mellor, G. L. and T. Yamada (1982). Development of a turbulence closure model for geophysical fluid problems. *Review of Geophysics* **20**(4): 851-875.

Molders, N. (2001). On the Uncertainty in Mesoscale Modeling Caused by Surface Parameters. *Meteorology and Atmospheric Physics* **76**: 119-141.

Moran, M. S., K. S. Humes and P. J. Pinter (1997). The scaling characteristics of remotely sensed variables for sparsely vegetated heterogeneous landscapes. *Journal of Hydrology* **190**: 337-362.

MRCC (2002). Historical Climate Summaries, Midwestern Regional Climate Center.

Nash, J. E. and J. V. Sutcliffe (1970). River flow forecasting through conceptual models 1. A discussion of principles. *Journal of Hydrology* **10**: 282-290.

Nicholls, M. E., A. S. Denning, L. Prihodko, P.-L. Vidale, I. Baker, K. J. Davis and P. S. Bakwin (2004). A Multiple-scale simulation of variations in atmospheric carbon dioxide using a coupled biosphere-atmospheric model.

Niyogi, D. S., S. Raman and K. Alapaty (1999). Uncertainty in the specification of surface characteristics, Part II: Hierarchy of interaction-explicit statistical analysis. *Boundary Layer Meteorology* **91**: 341-366.

Oechel, W., G. L. Vourlitis, J. Verfaille, T. Crawford, T. Brooks, E. Dumas, A. Hope, D. Stow, B. Boynton, V. Nosov and R. Zulueta (2000). A scaling approach for quantifying the net CO2 flux of the Kuparuk River Basin, Alaska. *Global Change Biology* **6**: 160-173.

O'Neill, R. V., R. H. Gardner and J. H. Carney (1982). Parameter constraints in a stream ecosystem model: incorporation of apriori information in monte carlo error analysis. *Ecological Modelling* **16**: 51-65.

Pacala, S. W., G. C. Hurtt, D. Baker, P. Peylin, R. A. Houghton, R. A. Birdsey, L. Heath, E. T. Sundquist, R. F. Stallard, P. Ciais, P. Moorcroft, J. P. Caspersen, E. Shevliakova, B. Moore, G. Kohlmaier, E. Holland, M. Gloor, M. E. Harmon, S. M. Fan, J. L. Sarmiento, C. L. Goodale, D. Schimel and C. B. Field (2001). Consistent land- and atmosphere-based US carbon sink estimates. *Science* **292**(5525): 2316-2320.

Pielke Jr., R. and R. T. Conant (2003). Best practices in prediction for decision-making: lessons from the atmospheric and earth sciences. *Ecology* **84**(6): 1351-1358.

Pielke, R. A., R. Avissar, M. Raupach, A. J. Dolman, X. B. Zeng and A. S. Denning (1998). Interactions between the atmosphere and terrestrial ecosystems: influence on weather and climate. *Global Change Biology* **4**(5): 461-475.

Pielke, R. A., W. R. Cotton, R. L. Walko, C. J. Tremback, W. A. Lyons, L. D. Grasso, M. E. Nicholls, M. D. Moran, D. A. Wesley, T. J. Lee and J. H. Copeland (1992). A Comprehensive Meteorological Modeling System - Rams. *Meteorology and Atmospheric Physics* **49**(1-4): 69-91.

Pielke, R. A., D. S. Schimel, T. J. Lee, T. Kittel and X. Zeng (1993). Atmosphere-terrestrial ecosystem interactions: implications for coupled modeling. *Ecological Modelling*(67): 5-18.

Pielke Sr., R. A., R. L. Walko, L. Staeyart, P.-L. Vidale, G. Liston, W. A. Lyons and T. N. Chase (1999). The influence of anthropogenic landscape changes on weather in southern Florida. *Monthly Weather Review* **127**: 1663-1673.

Pitman, A., R. Pielke Sr., R. Avissar, M. Claussen, J. Gash and H. Dolman (1999). The role of the land surface in weather and climate: does the land surface matter? Global Change Newsletter**:** 4-11.

Potter, C. S., S. S. Wang, N. T. Nikolov, A. D. McGuire, J. Liu, A. W. King, J. S. Kimball, R. F. Grant, S. E. Frolking, J. S. Clein, J. M. Chen and J. S. Amthor (2001). Comparison of boreal ecosystem model sensitivity to variability in climate and forest site parameters. *Journal of Geophysical Research-Atmospheres* **106**(D24): 33671-33687.

Prince, S. D., Y. H. Kerr, J. P. Goutorbe, T. Lebel, A. Tinga, P. Bessemoulin, J. Brouwer, A. J. Dolman, E. T. Engman, J. H. C. Gash, M. Hoepffner, P. Kabat, B. Monteny, F. Said, P. Sellers and J. Wallace (1995). Geographical, Biological and Remote-Sensing Aspects of the Hydrologic Atmospheric Pilot Experiment in the Sahel (Hapex-Sahel). *Remote Sensing of Environment* **51**(1): 215-234.

Rahman, A. F., J. A. Gamon, D. A. Fuentes, D. A. Roberts and D. Prentiss (2001). Modeling spatially distributed ecosystem flux of boreal forest using hyperspectral indices from AVIRIS imagery. *Journal of Geophysical Research* **106**(D24): 33,579-33,591.

Raich, J. W. and K. J. Nadelhoffer (1989). Belowground carbon allocation in forest ecosystems: Global trends. *Ecology* **70**(5): 1346-1354.

Raich, J. W., E. B. Rastetter, J. M. Melillo, D. W. Kicklighter, P. A. Steudler, B. J. Peterson, A. L. Grace, B. Moore III and C. J. Vorosmarty (1991). Potential net primary productivity in South America: Application of a global model. *Ecological Applications* **1**(4): 399-429.

Randall, D. A., D. A. Dazlich, C. Zhang, A. S. Denning, P. J. Sellers, C. J. Tucker, L. Bounoua, S. O. Los, C. O. Justice and I. Fung (1996). A revised land surface parameterization (SiB2) for GCMs .3. The greening of the Colorado State University general circulation model. *Journal of Climate* **9**(4): 738-763.

Rayner, P. J., I. G. Enting, R. J. Francey and R. Langenfelds (1999). Reconstructing the recent carbon cycle from atmospheric CO2, d13C and O2/N2 observations. *Tellus* **51B**: 213-232.

Rodenbeck, C., S. Houweling, M. Gloor and M. Heimann (2003). CO2 flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics* **3**: 1919-1964.

Romanowicz, R., K. J. Beven and J. A. Tawn (1994). Evaluation of predictive uncertainty in nonlinear hydrological models using a Bayesian approach. Statistics for the Environment 2. V. Barnett and K. F. Turkman. Chichester, John Wiley and Sons: 297-318.

Roy, S. B. and R. Avissar (2002). Impact of land use/land cover change on regional hydrometeorology in Amazonia. *Journal of Geophysical Research-Atmospheres* **107**(D20).

Roy, S. B., C. P. Weaver, D. S. Nolan and R. Avissar (2003). A preferred scale for landscape forced mesoscale circulations? *Journal of Geophysical Research-Atmospheres* **108**(D22).

Saltelli, A. (1999). Sensitivity analysis: Could better methods be used? *Journal of Geophysical Research* **104**(D3): 3789-3793.

Schimel, D., J. Melillo, H. Q. Tian, A. D. McGuire, D. Kicklighter, T. Kittel, N. Rosenbloom, S. Running, P. Thornton, D. Ojima, W. Parton, R. Kelly, M. Sykes, R. Neilson and B. Rizzo (2000). Contribution of increasing CO2 and climate to carbon storage by ecosystems in the United States. *Science* **287**(5460): 2004-2006.

Schimel, D. S. (1995). Terrestrial ecosystems and the carbon cycle. *Global Change Biology* **1**: 77-91.

Schlesinger, W. H. (1997). Biogeochemistry: an Analysis of Global Change. San Diego, Academic Press.

Schulz, K., K. J. Beven and B. Huwe (1999). Equifinality and the problem of robust calibration in Nitrogen budget simulations. *Soil Science Society of America Journal* **63**: 1934-1941.

Schulz, K., A. Jarvis, K. J. Beven and H. Soegaard (2001). The predictive uncertainty of land surface fluxes in response to increasing ambient carbon dioxide. *Journal of Climate* **14**: 2551-2562.

Segal, M., R. Avissar, M. C. McCumber and R. A. Pielke (1988). Evaluation of vegetation effects on the generation and modification of mesoscale circulations. *Journal of the Atmospheric Sciences* **45**(16): 2268-2292.

Sellers, P., F. Hall, H. Margolis, B. Kelly, D. Baldocchi, G. Denhartog, J. Cihlar, M. G. Ryan, B. Goodison, P. Crill, K. J. Ranson, D. Lettenmaier and D. E. Wickland (1995a). The Boreal Ecosystem-Atmosphere Study (Boreas) - an overview and early results from the 1994 field year. *Bulletin of the American Meteorological Society* **76**(9): 1549-1577.

Sellers, P., M. D. Heiser, F. Hall, S. J. Goetz, D. E. Strebel, S. Verma, R. L. Desjardins, P. M. Scheupp and J. I. MacPherson (1995b). Effects of spatial variability in topography, vegetation cover and soil moisture on area-averaged surface fluxes: A case study from the FIFE 1989 data. *Journal of Geophysical Research* **100**(D12): 25,607-25,629.

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing* **6**(8): 1335-1371.

Sellers, P. J. (1987). Canopy reflectance, photosynthesis and transpiration. II. The role of biophysics in the linearity of their independence. *Remote Sensing of Environment* **21**: 143-183.

Sellers, P. J., J. A. Berry, G. J. Collatz, C. B. Field and F. G. Hall (1992a). Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of Environment* **42**: 187-216.

Sellers, P. J., R. E. Dickinson, D. A. Randall, A. K. Betts, F. G. Hall, J. A. Berry, G. J. Collatz, A. S. Denning, H. A. Mooney, C. A. Nobre, N. Sato, C. B. Field and A. HendersonSellers (1997a). Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* **275**(5299): 502-509.

Sellers, P. J., F. G. Hall, G. Asrar, D. E. Strebel and R. E. Murphy (1992b). An Overview of the 1st International Satellite Land Surface Climatology Project (Islscp) Field Experiment (Fife). *Journal of Geophysical Research-Atmospheres* **97**(D17): 18345-18371.

Sellers, P. J., M. D. Heiser, F. G. Hall, S. B. Verma, R. L. Desjardins, P. M. Schuepp and J. I. MacPherson (1997b). The impact of using area-averaged land surface properties - Topography, vegetation condition, soil wetness - In calculations of intermediate scale (approximately 10 km(2)) surface-atmosphere heat and moisture fluxes. *Journal of Hydrology* **190**(3-4): 269-301.

Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz and D. A. Randall (1996a). A revised land surface parameterization (SiB2) for atmospheric GCMs .2. The generation of global fields of terrestrial biophysical parameters from satellite data. *Journal of Climate* **9**(4): 706-737.

Sellers, P. J., D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, D. A. Dazlich, C. Zhang, G. D. Collelo and L. Bounoua (1996b). A revised land surface parameterization (SiB2) for atmospheric GCMs .1. Model formulation. *Journal of Climate* **9**(4): 676-705.

Smagorinsky, J. S. (1963). General circulation experiments with the primitive equations. Part I: The basic experiment. *Monthly Weather Review* **91**: 96-164.

Sokal, R. R. and F. J. Rohlf (1981). Biometry. New York, W.H. Freeman and Company.

Sokal, R. R. and F. J. Rohlf (1995). Biometry. New York, W.H. Freeman and Company.

Spear, R. C., T. M. Grieb and N. Shang (1994). Parameter uncertainty and interaction in complex environmental models. *Water Resources Research* **30**(11): 3159-3169.

Staff, S. S. (1994). State Soil Geographic Database (STATSGO). Fort Worth, U.S. Department of Agriculture, National Conservation Service.

Stohlgren, T. J., T. N. Chase, R. A. Pielke Sr., T. Kittel and J. S. Baron (1998). Evidence that local land use practices influence regional climate, vegetation, and stream flow patterns in adjacent natural areas. *Global Change Biology* **5**(4): 495-504.

Tans, P. P., I. Y. Fung and T. Takahashi (1990). Observational constraints on the global atmospheric CO_2 budget. *Science* **247**: 1431-1438.

Tans, P. P. and J. W. C. White (1998). In balance, with a little help from the plants. *Science* **281**: 183-184.

Teillet, P. M., E. L. Saleous, M. C. Hansen, J. C. Eidenshink, C. O. Justice and J. R. G. Townshend (2000). An evaluation of the global 1 Km AVHRR land dataset. *International Journal of Remote Sensing* **21**(10): 1987-2021.

USDAFS (2001). Chequamegon-Nicolet National Forest Area History, USDA Forest Service Chequamegon-Nicolet National Forest.

Valentini, R., G. Matteucci, A. J. Dolman, E. D. Schulze, C. Rebmann, E. J. Moors, A. Granier, P. Gross, N. O. Jensen, K. Pilegaard, A. Lindroth, A. Grelle, C. Bernhofer, T. Grunwald, M. Aubinet, R. Ceulemans, A. S. Kowalski, T. Vesala, U. Rannik, P. Berbigier, D. Loustau, J. Guomundsson, H. Thorgeirsson, A. Ibrom, K. Morgenstern, R. Clement, J. Moncrieff, L. Montagnani, S. Minerbi and P. G. Jarvis (2000). Respiration as the main determinant of carbon balance in European forests. *Nature* **404**(6780): 861-865.

Ver Hoef, J. M. (1996). Parametric empirical Bayes methods for ecological applications. *Ecological Applications* **6**(4): 1047-1055.

Veroustraete, F., J. Patyn and R. B. Myneni (1996). Estimating net ecosystem exchange of carbon using the normalized difference vegetation index and an ecosystem model. *Remote Sensing of Environment* **58**: 115-130.

Vourlitis, G. L., W. Oechel, A. Hope, D. Stow, B. Boynton, J. Verfaille, R. Zulueta and S. J. Hastings (2000). Physiological models for scaling plot measurements of CO2 flux across an arctic tundra landscape. *Ecological Applications* **10**(1): 60-72.

Walko, R. L., L. E. Band, J. Baron, T. G. F. Kittel, R. Lammers, T. J. Lee, D. Ojima, R. A. Pielke, C. Taylor, C. Tague, C. J. Tremback and P. L. Vidale (2000). Coupled atmosphere-biophysics-hydrology models for environmental modeling. *Journal of Applied Meteorology* **39**(6): 931-944.

Walko, R. L., C. J. Tremback, R. Pielke and W. R. Cotton (1995). An interactive nesting algorithm for stretched grids and variable nesting ratios. *Journal of Applied Meteorology* **34**: 994-999.

Weaver, C. P. and R. Avissar (2001). Atmospheric disturbances caused by human modification of the landscape. *Bulletin of the American Meteorological Society* **82**(2): 269-281.

Weaver, C. P., S. B. Roy and R. Avissar (2002). Sensitivity of simulated mesoscale atmospheric circulations resulting from landscape heterogeneity to aspects of model configuration. *Journal of Geophysical Research* **107**(D20): 8041.

WiDNR Wisconsin Department of Natural Resources (1998). WISCLAND Land Cover (WLCGW930). Madison, Wisconsin, WiDNR.

Wikle, C. K. (2003). Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* **84**(6): 1382-1394.

Wilson, K., A. Goldstein, E. Falge, M. Aubinet, D. Baldocchi, P. Berbigier, C. Bernhofer, R. Ceulemans, H. Dolman, C. Field, A. Grelle, A. Ibrom, B. E. Law, A. Kowalski, T. Meyers, J. Moncrieff, R. Monson, W. Oechel, J. Tenhunen, R. Valentini and S. Verma (2002). Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology* **113**(1-4): 223-243.

Wofsy, S. C. and R. C. Harriss (2002). The North American Carbon Program. Report of the NACP Committee of the U.S. Interagency Carbon Cycle Science Program. Washington, D.C., U.S. Global Change Research Program: 62.

Wood, E. F. (1995). Scaling behavior of hydrological fluxes and variables: empirical studies using a hydrological model and remote sensing data. *Hydrological Processes* **9**: 331-346.

Xia, Y., A. J. Pitman, H. V. Gupta, M. Leplastrier, A. Henderson-Sellers and L. Bastidas (2002). Calibrating a land surface model of varying complexity using multicriteria methods and the Cabauw dataset. *Journal of Hydrometeorology* **3**.

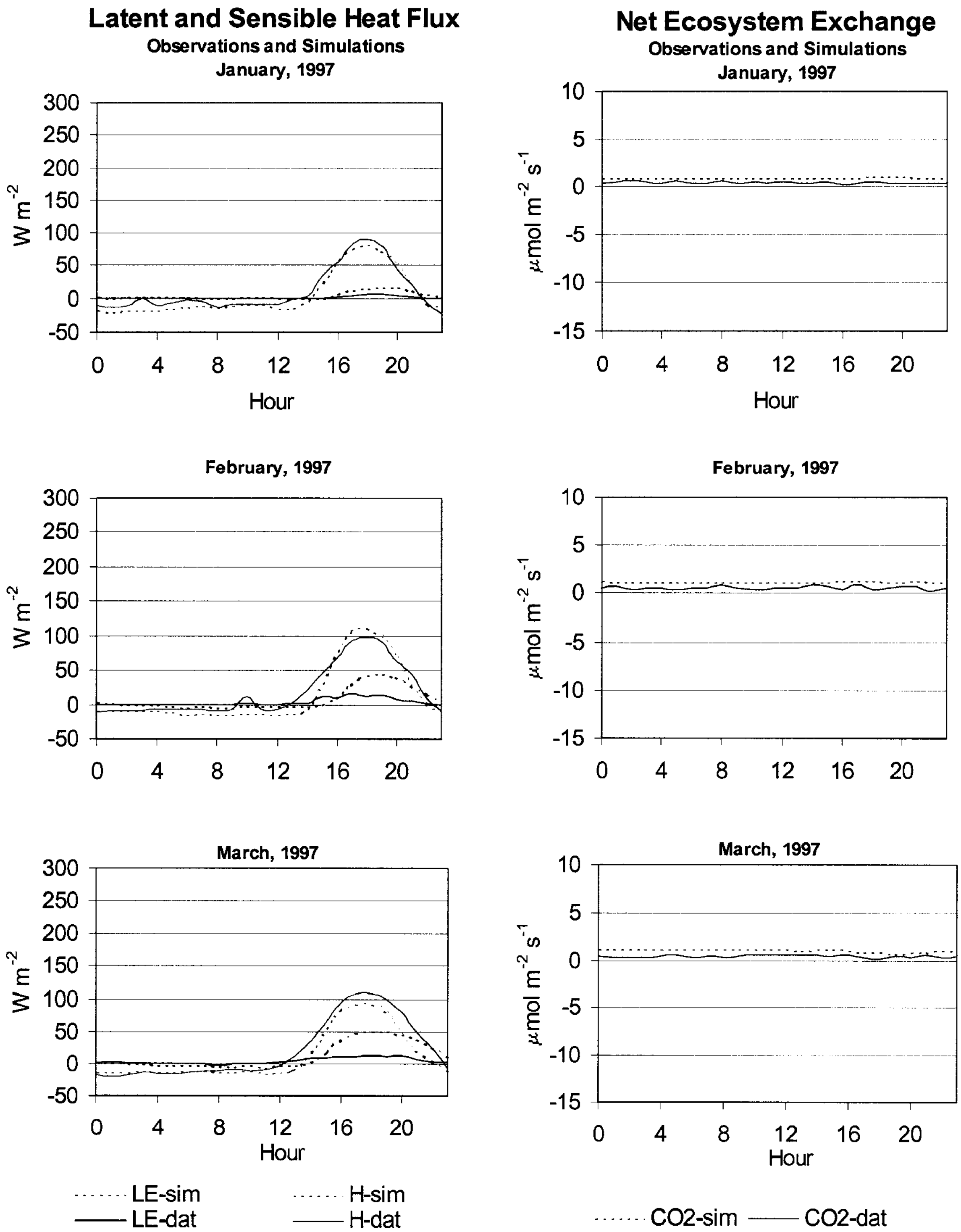
Yates, D. N., F. Chen and H. Nagai (2003). Land surface heterogeneity in the Cooperative Atmosphere Surface Exchange Study (CASES-97). Part II: Analysis of spatial heterogeneity and its scaling. *Journal of Hydrometeorology* **4**: 219-234.

Yi, C. X., K. J. Davis, B. W. Berger and P. S. Bakwin (2001). Long-term observations of the dynamics of the continental planetary boundary layer. *Journal of the Atmospheric Sciences* **58**(10): 1288-1299.

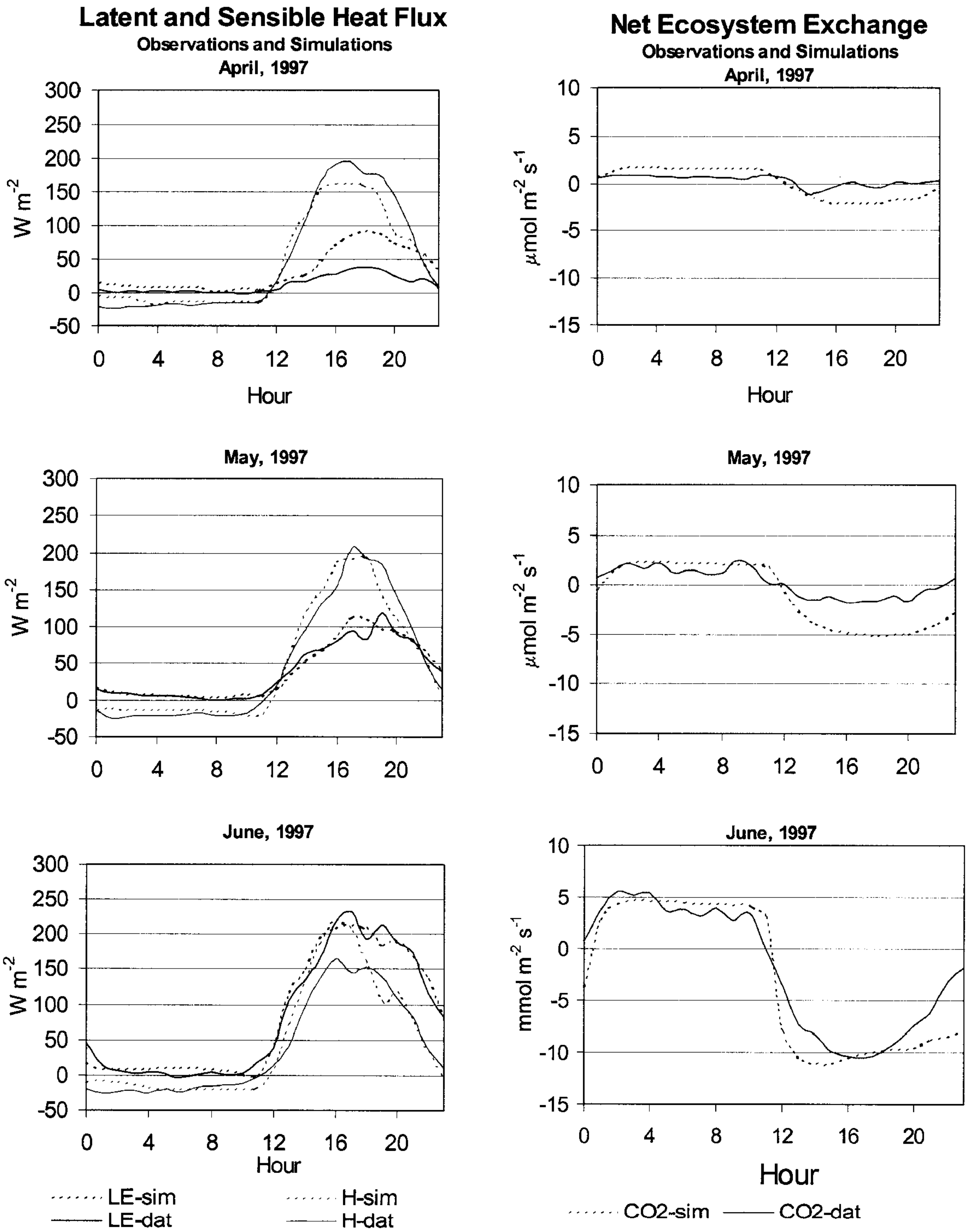
APPENDIX 1

Monthly Diurnal Mean Plots

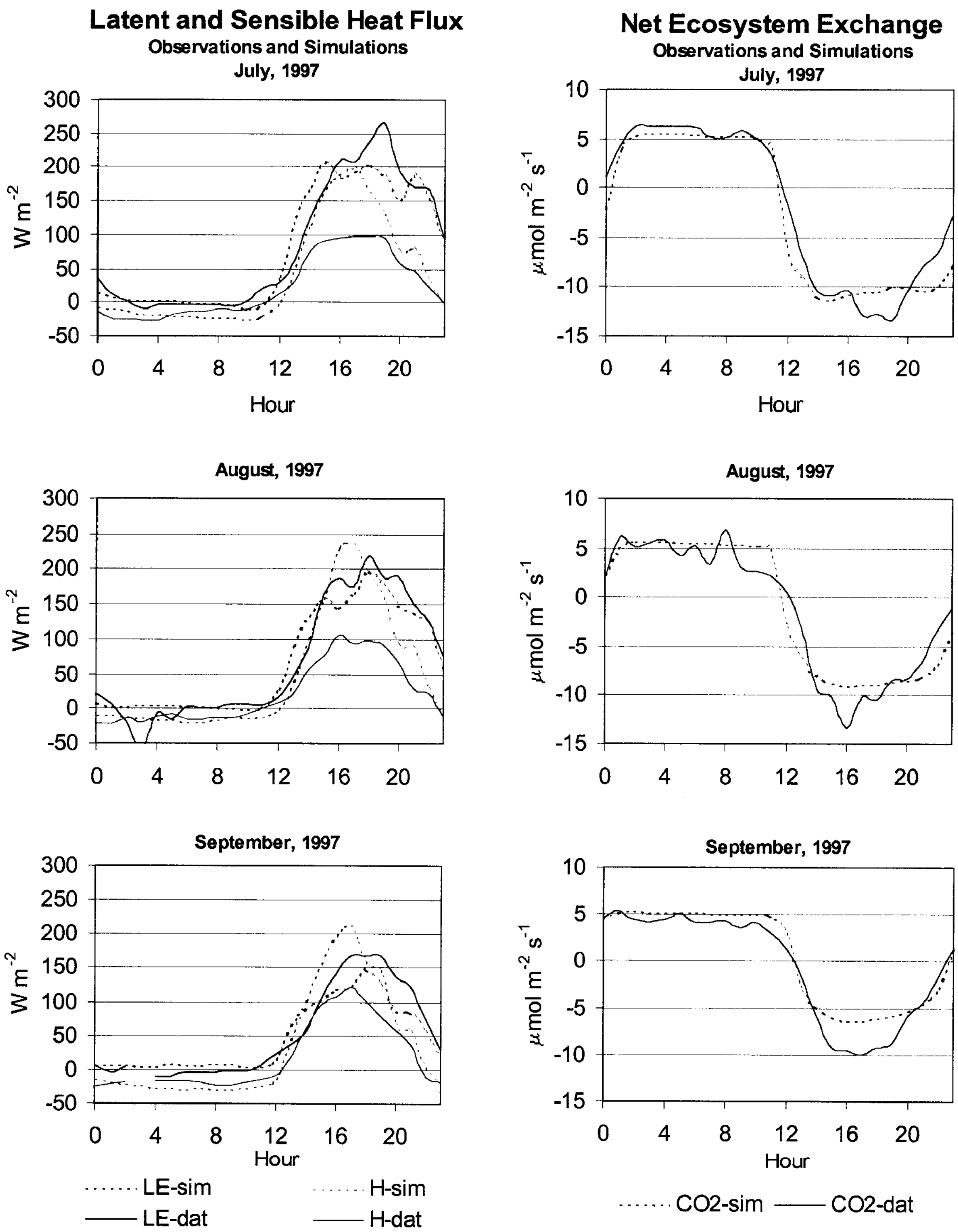
Each plot in this appendix shows a monthly mean diurnal composite of both observed and simulated fluxes. Solid lines represent observations and dashed lines represent simulations. Red indicates sensible heat flux, blue indicates latent heat flux and green indicates the net ecosystem exchange of Carbon. All twelve months of 1997 are shown.

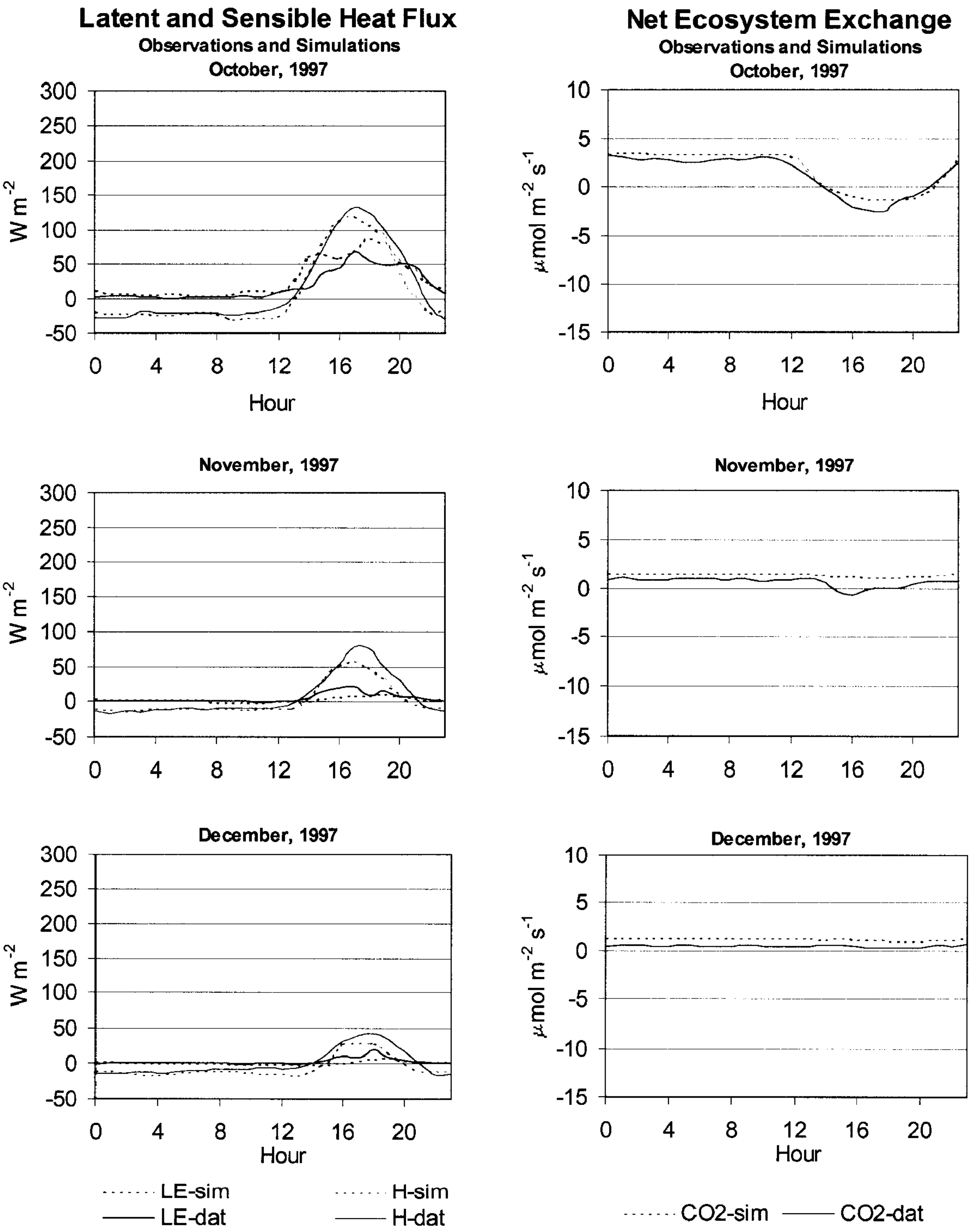


Latent and Sensible Heat Flux
Observations and Simulations
January, 1997
February, 1997
March, 1997
W m-2
300
250
200
150
100
50
0
-50
0
4
8
12
16
20
Hour
Net Ecosystem Exchange
Observations and Simulations
January, 1997
February, 1997
March, 1997
µmol m-2 s-1
10
5
0
-5
-10
-15
LE-sim
LE-dat
H-sim
H-dat
CO2-sim
CO2-dat



Latent and Sensible Heat Flux
Observations and Simulations
April, 1997
Net Ecosystem Exchange
Observations and Simulations
April, 1997
May, 1997
June, 1997
W m-2
μmol m-2 s-1
mmol m-2 s-1
Hour
LE-sim
LE-dat
H-sim
H-dat
CO2-sim
CO2-dat


Latent and Sensible Heat Flux
Observations and Simulations
July, 1997
Net Ecosystem Exchange
Observations and Simulations
July, 1997
W m-2
μmol m-2 s-1
Hour
August, 1997
August, 1997
September, 1997
September, 1997
LE-sim
H-sim
LE-dat
H-dat
CO2-sim
CO2-dat


Latent and Sensible Heat Flux
Observations and Simulations
October, 1997
Net Ecosystem Exchange
Observations and Simulations
October, 1997
W m-2
µmol m-2 s-1
Hour
November, 1997
November, 1997
December, 1997
December, 1997
Hour
Hour
LE-sim
LE-dat
H-sim
H-dat
CO2-sim
CO2-dat

APPENDIX 2

Equations for Time Varying Parameters

Listed below are some of the equations used to calculate time-varying parameters for SiB2.5 from the normalized difference vegetation index (NDVI). These equations can be found in Sellers et al (Sellers *et al.* 1996) and Los et al (Los *et al.* 2000) but are reproduced here to aid in the interpretation of results. Also listed below are definitions of the parameters used in the equations. Parameters in **bold** are varied in the sensitivity analysis and are essential to the equations that produce the time varying parameters, which are displayed in red.

Definitions:

NDVI	= normalized difference vegetation index; NIR-R/NIR+R
NIR	= near infrared reflectance
R	= red reflectance
SR	= simple ratio; NIR/R
f_{PAR}	= fraction of absorbed photosynthetically active radiation
SRf_{PAR}	= f_{PAR} calculated from the simple ratio
NDVIf_{PAR}	= f_{PAR} calculated from NDVI
ND98	= 98^{th} percentile NDVI by biome
ND02	= 2^{nd} percentile NDVI by biome
f_{PAR}max	= maximum possible f_{PAR} (0.95)
f_{PAR}min	= minimum possible f_{PAR} (0.01)
fvcover	= fraction of surface covered by vegetation
Green	= green fraction of LAI
ZLT	= leaf area index
LAIg	= green leaf area index at current time
LAId	= dead leaf area index at current time
stem	= stem area index
LTmax	= maximum possible leaf area index by biome
f_{PAR}m	= f_{PAR} from the previous time step
scatp	= canopy transmittance and reflectance coefficient with respect to PAR
PARk	= mean canopy absorption optical depth with respect to PAR
LTran(1,1)	= transmittance of green vegetation
LTran(1,2)	= transmittance of brown vegetation
LRef(1,1)	= reflectance of green vegetation
LRef(1,2)	= reflectance of brown vegetation
gmudmu	= time mean projected leaf area normal to the sun
aa	= minimum possible leaf area projection versus cosine of the sun angle
bb	= slope of leaf area projection versus cosine of the sun angle
fb	= PAR weighted mean cosine of the sun angle
chil	= leaf angle distribution factor
topint	= total flux onto the canopy adjusted for sun angle in a 24 hour period
botint	= total PAR flux onto the canopy in a 24 hour period

$$f_{PAR} = 0.5*(SRf_{PAR} + NDVIf_{PAR}) \quad \text{(A.1.1)}$$

$$SRf_{PAR} = (SR - SRmin)*(f_{PAR}max - f_{PAR}min)/(SRmax - SRmin) + f_{PAR}min \quad \text{(A.1.2)}$$

$$SR = \frac{(1+NDVI)}{(1-NDVI)} \quad \text{(A.1.3)}$$

$$SRmax = \frac{1+\mathbf{ND98}}{1-\mathbf{ND98}} \quad \text{(A.1.4)}$$

$$SRmin = \frac{1+\mathbf{ND02}}{1-\mathbf{ND02}} \quad \text{(A.1.5)}$$

$$NDVIf_{PAR} = (NDVI - \mathbf{ND02})*(f_{PAR}max - f_{PAR}min)/(\mathbf{ND98} - \mathbf{ND02}) + f_{PAR}min \quad \text{(A.1.6)}$$

$$ZLT = (LAIg + LAId + \mathbf{stem})*\mathbf{fvcover} \quad \text{(A.1.7)}$$

$$LAIg = a\log(1 - f_{PAR} / \mathbf{fvcover})*\mathbf{LTmax} / a\log(1 - f_{PAR}max) \quad \text{(A.1.8)}$$

If LAIg is greater than or equal to LAIg of the previous time step:

$$LAId = 0.0001 \quad \text{(A.1.9)}$$

If LAIg is less than LAIg of the previous time step:

$$LAId = 0.5*(LAIg(previous) - LAIg(current)) \quad \text{(A.1.10)}$$

$$f_{PAR}new = \mathbf{fvcover}*(1 - \exp(PARk*ZLT / \mathbf{fvcover})) \quad \text{(A.1.11)}$$

$$PARk = \sqrt{(1 - scatp)}*gmudmu \quad \text{(A.1.12)}$$

$$scatp = Green*(\mathbf{LTran(1,1)} + \mathbf{LRef(1,1)}) + (1 - Green)*(\mathbf{LTran(1,2)} + \mathbf{LRef(1,2)}) \quad \text{(A.1.13)}$$

$$Green = LAIg/(LAIg + LAId + \mathbf{stem}) \quad \text{(A.1.14)}$$

$$gmudmu = (aa + bb*fb)/fb \quad \text{(A.1.15)}$$

$$aa = 0.5 - 0.633*\mathbf{chil} - 0.33*\mathbf{chil}^2 \quad \text{(A.1.16)}$$

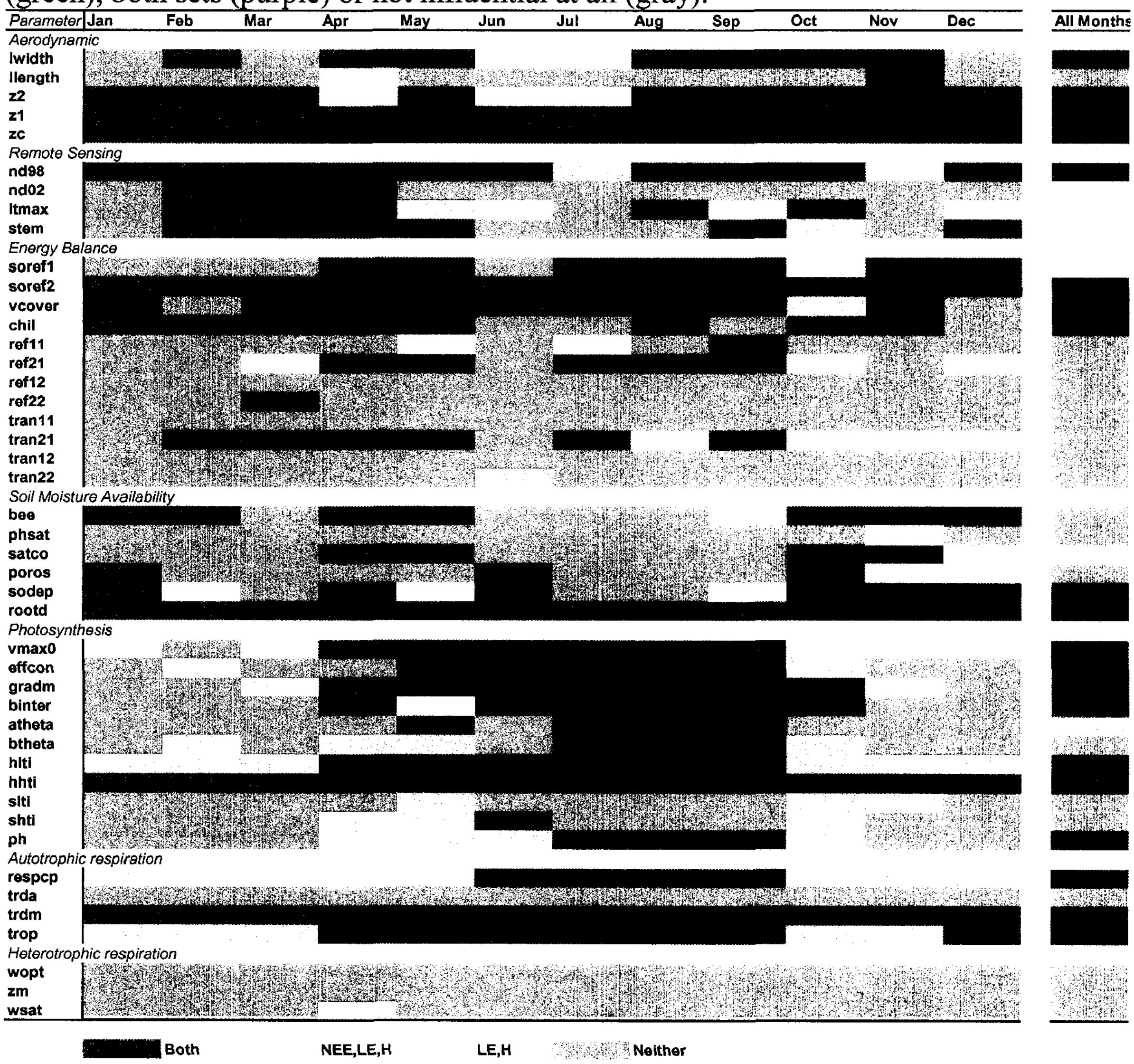
$$bb = 0.877(1 - 2*aa) \quad \text{(A.1.17)}$$

$$fb = topint / botint \quad \text{(A.1.18)}$$

APPENDIX 3

Parameter Influence for the computed Multi-Objective Criteria

Table A3.1 Influential (p = 0.01) and non-influential model parameters as determined by the Kolmogorov-Smirnov test statistic. Parameters are arranged by usage in the model and shaded to reflect whether the parameter was found influential in the prediction of latent and sensible heat (yellow), latent and sensible heat plus net ecosystem exchange (green), both sets (purple) or not influential at all (gray).

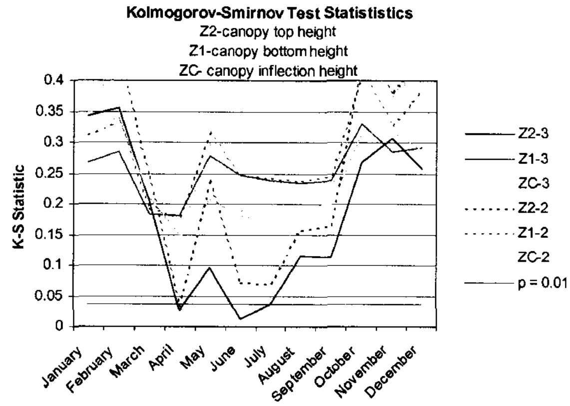


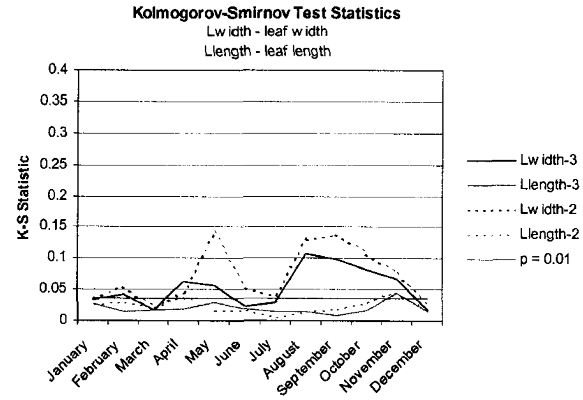
APPENDIX 4

Kolomogorov-Smirnov Test Statistic Plots

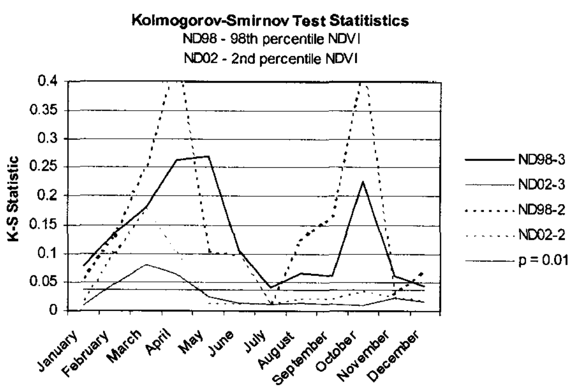
Each plot in this appendix shows two or more parameters, the solid lines represent the NEE,LE,H criterion denoted by '-3' in the legend and the dashed lines represent the LE,H criterion denoted by '-2' in the legend. The red line indicates significance at the p=0.01 level. Points that fall above this line are influential and the height of the point indicates the degree of sensitivity of the model to the specification of that parameter. All of the plots are in the same format and have the same scale.

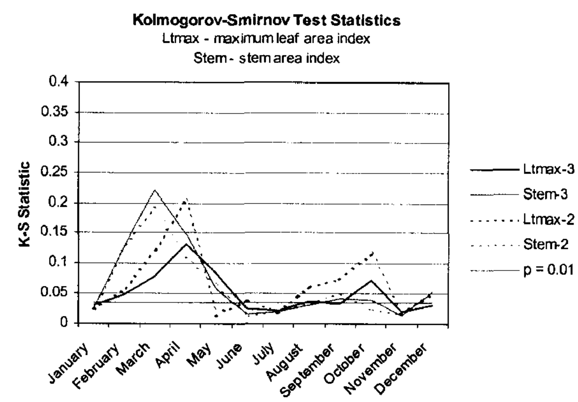
Aerodynamic Parameters



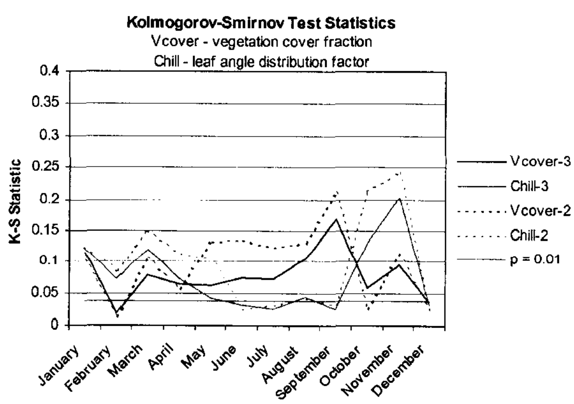


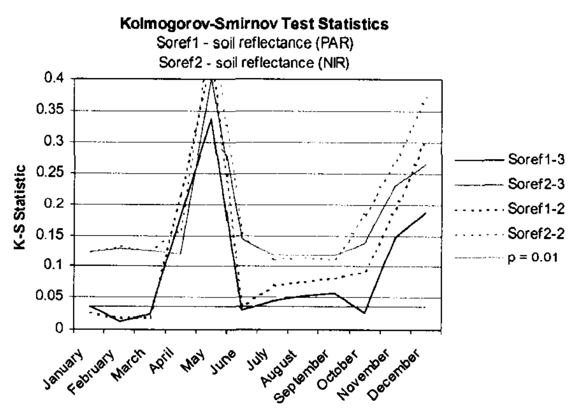
Remote Sensing Parameters



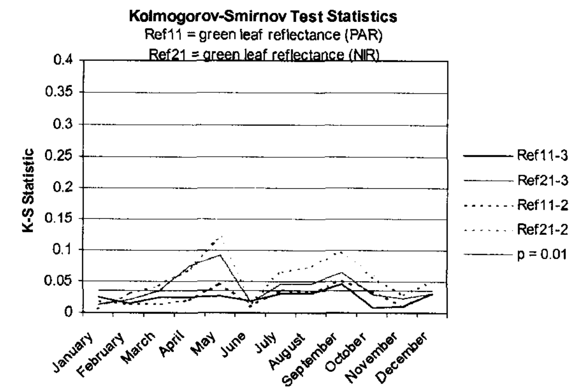


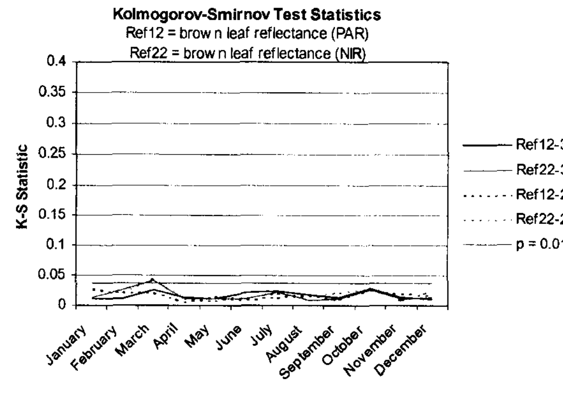
Energy Balance Parameters

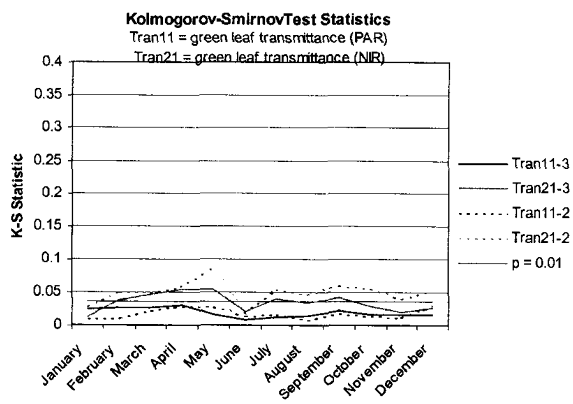


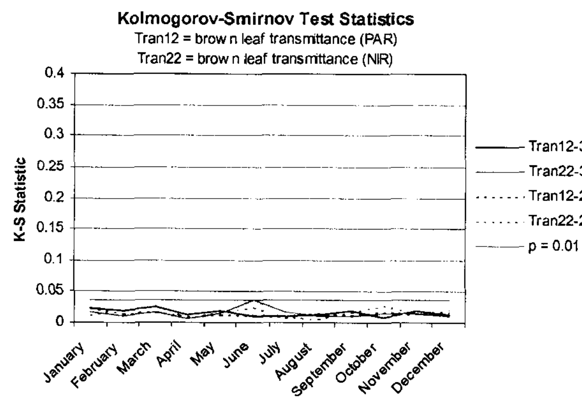


Energy Balance Parameters cont'd

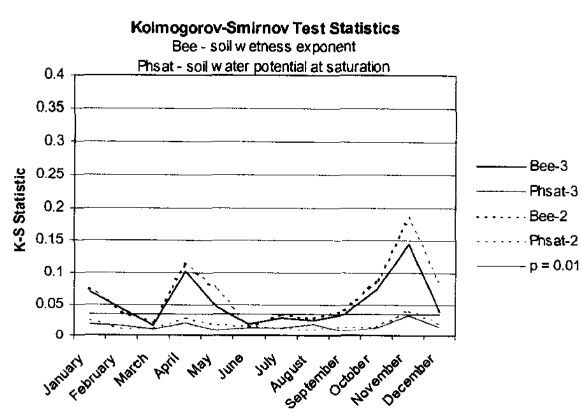


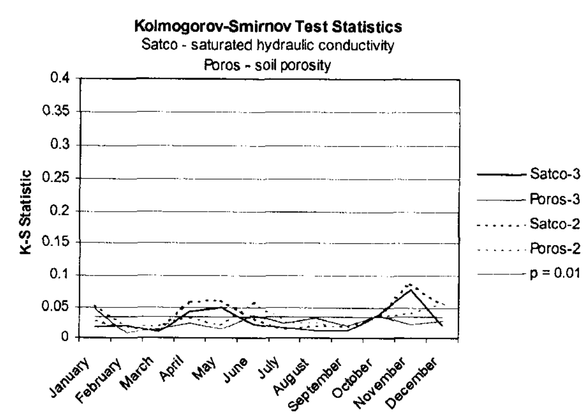




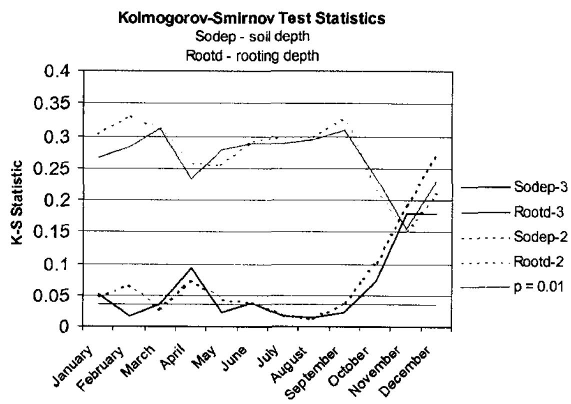


Soil moisture availability parameters

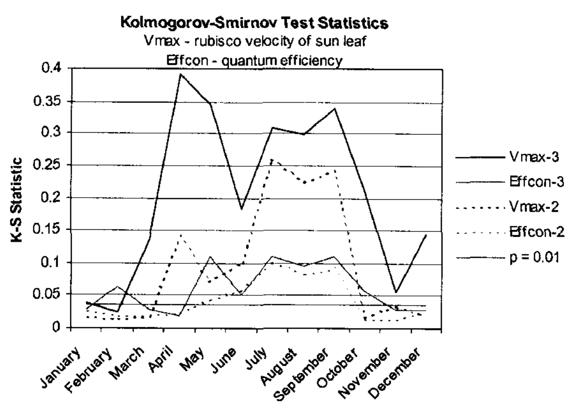


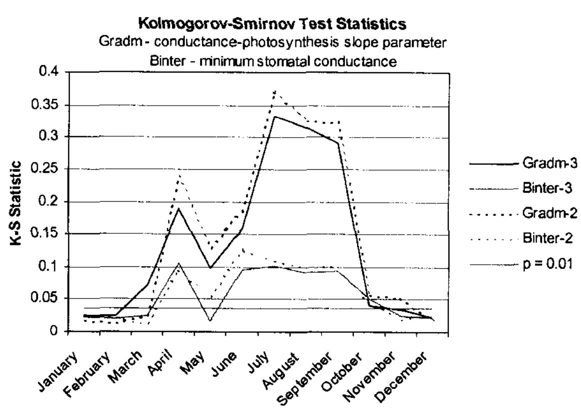


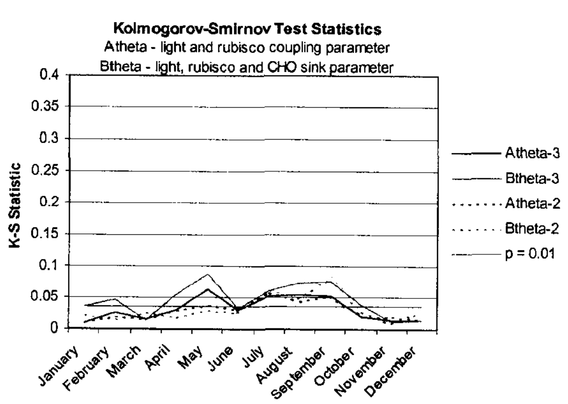
Soil moisture availability parameters cont'd

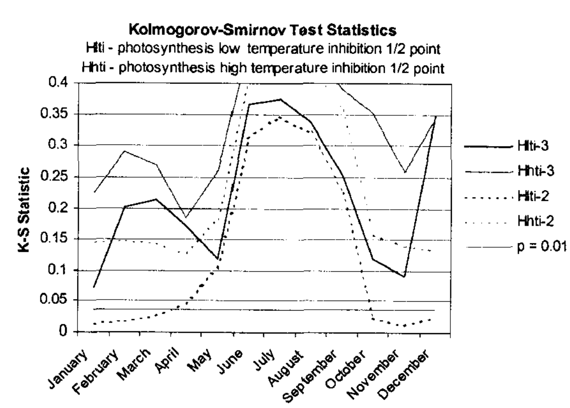


Photosynthesis Parameters

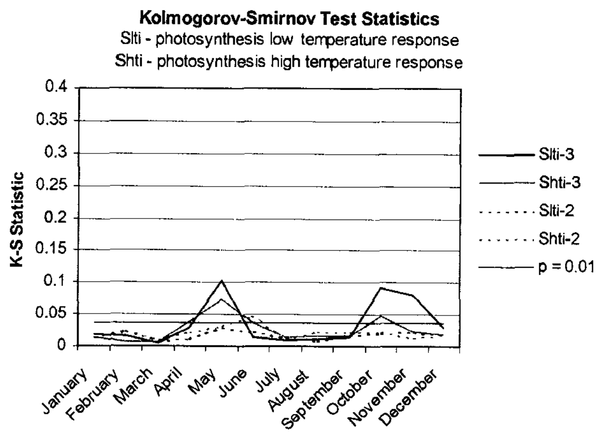




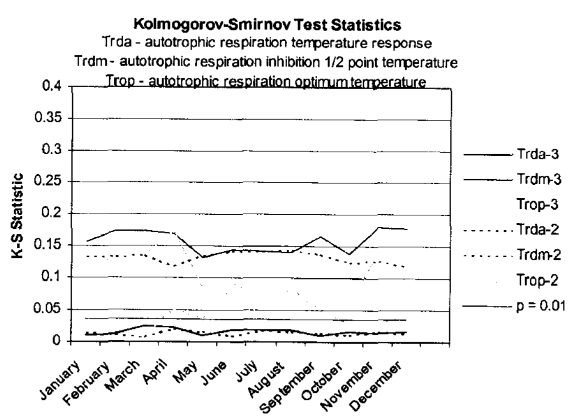


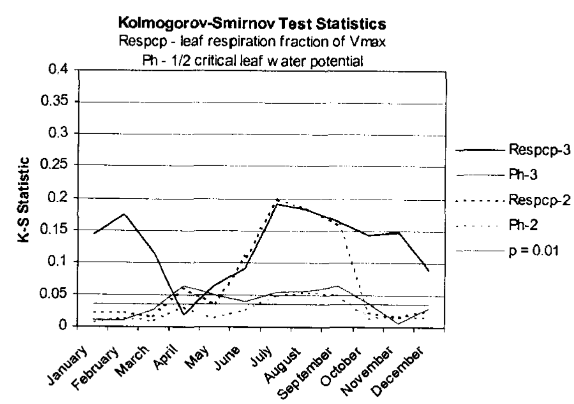


Photosynthesis Parameters cont'd



Autotrophic Respiration Parameters





Heterotrophic Respiration Parameters