PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Sanjiv Kumar

Entitled
Land Use Land Cover Change and Atmospheric Feedback: Impact on Regional Water
Resources

For the degree of   Doctor of Philosophy

Is approved by the final examining committee:

Venkatesh M. Merwade
_____
            Chair
Laura C. Bowling

Rao S. Govindaraju

P. S. C. Rao

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Venkatesh M. Merwade
                               _____

                               _____

Approved by: Garrett D. Jeong   04/25/2011
            Head of the Graduate Program   Date

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

Land Use Land Cover Change and Atmospheric Feedback: Impact on Regional Water
Resources

For the degree of   Doctor of Philosophy

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Sanjiv Kumar
_____
Printed Name and Signature of Candidate

04/22/2011
_____
Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

LAND USE LAND COVER CHANGE AND ATMOSPHERIC FEEDBACK: IMPACT ON
REGIONAL WATER RESOURCES

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Sanjiv Kumar

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2011

Purdue University

West Lafayette, Indiana

To my parents and wife

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Kumar, Sanjiv. Ph.D., Purdue University, May 2011. Land Use Land Cover Change and Atmospheric Feedback: Impact on Regional Water Resources. Major Professor: Venkatesh M. Merwade

Land Use and Land Cover (LULC) change, such as conversion of natural vegetation into agricultural land and urbanization, is a major global change phenomenon. Between 1700 and 2000, the global extent of natural vegetation has decreased by 45%, and agricultural land area has increased five fold. LULC change impacts hydrology by changing regional climate (temperature and precipitation) and land surface hydrologic responses (evapotranspiration, runoff, and ground water recharge). Evapotranspiration (ET) is a major pathway through which the land surface interacts with the atmosphere, and it is a major component (60-65% of total precipitation) of the global hydrologic cycle. The three objectives of this study are to: (1) investigate impacts of regional scale LULC change on the regional hydroclimatology, and compare impacts of LULC change with impacts of climate change arising from elevated green house gas emissions, (2) evaluate uncertainties in reanalysis and climate model ET outputs using AmeriFlux observations and a basin scale water and energy balance studies, and (3) quantify contributions of major driving forces for LULC change in the United States.

Large scale drainage of wetlands was carried out in the latter half of the 19th century and early half of the 20th century to bring swamp/marshy land of the Midwestern United States (pre-settlement landscape) into intensive agricultural production (Corn Belt of USA). Impacts of wetland drainage (LULC change) on hydroclimatology of the Midwestern United States are compared with impacts of climate change using a coupled land-atmosphere global climate model (CCSM3). The wetland drainage data are obtained from United States Census reports. Results from this study suggest that impacts of wetland drainage can be of comparable magnitude to impacts of climate change attributed to greenhouse gas emissions.

The Community Land Model (CLM) and North American Regional Reanalysis (NARR) outputs are evaluated using AmeriFlux observations, PRISM precipitation and temperature data, and USGS streamflow observations in the Mississippi River Basin. Based on averages over 11 AmeriFlux sites, NARR shows higher biases (59%) in ET compared to CLM (11%). Issues related to point scale observations versus climate model grid cell outputs, and model parametrization differences between CLM and NARR are also investigated in this study.

The land-cover change history of the United States is investigated to determine major driving/governing forces. County level cropland and population data from 1850 to 2000 (per decade), and high resolution topography, climate, and biophysical suitability data are used. Results from this study suggest that the spatial distribution of cropland was governed by population distribution during the 19th century, and biophysical suitability (for cropland) during the 20th century. The major influence of biophysical suitability is expected to continue in the near future landscape of the United States.

CHAPTER1. INTRODUCTION

1.1 <u>Background and Motivation</u>

Land Use Land Cover (LULC) change is an important phenomenon that has occurred globally over the past several centuries. Between 1700 and 2000, the global extent of natural vegetation has decreased by 45%, and agricultural land area has increased five fold (Pongratz et al., 2008, Scanlon et al., 2007). The different forms of LULC change, such as deforestation, agricultural intensification, urbanization, and reforestation continue to take place in different parts of the world (Lambin et al., 2001; Mustard et al., 2004). LULC change is a regionally significant phenomenon on decade to century time scales, and the rate and intensity of LULC change have changed with time and also from one region to another. For example, a majority of forested land-cover in the Northeastern United States was converted to agricultural land-cover during the 18th and 19th centuries (Williams, 1989); whereas, major LULC change in the Midwestern United States took place between 1850 and 1950 (Whitney, 1994).

The LULC change impacts hydrology through: (1) changes in land surface hydrologic response characteristics such as infiltration, runoff, and evapotranspiration, and (2) changes in regional climate (Bonan, 1997, 1999; Pielke, 2005). The conversion of natural vegetation into agricultural land decreases evapotranspiration, and hence increases fresh water availability (Gordon et al., 2003). Urbanization increases surface runoff and decreases groundwater recharge (Yang et al., 2009). LULC change also affects surface radiation and energy budgets through changes in surface albedo (ratio of reflected to incident solar radiation), Bowen ratio (ratio of sensible to latent heat flux) and roughness height (Bonan, 1999). The surface energy fluxes (sensible and latent heat) are major determinants of local and regional climate (Bonan, 2008, p. 192-203).

Evapotranspiration (ET) is a major pathway through which the land surface interacts with the atmosphere, and it is the major component of the global hydrologic cycle (60-65% of total precipitation, Postel et al., 1996). The energy and water fluxes within a hydrologic

2

boundary/basin are linked through ET. Despite the importance of ET, relatively few reliable observations of ET are available compared to runoff. The availability of tower flux data from FLUXNET sites in the recent decades has provided a new opportunity to improve our understanding of land surface and atmospheric interaction (Baldocchi et al., 2001).

A number of studies have found significant impacts of LULC changes on regional climate (Bonan, 1997, 1999; Li and Molders 2008; Findell et al., 2009). At the global scale, the opposing signs of regional hydroclimatic changes tend to nullify each other (e.g. making some regions cooler and other regions warmer), making their global average impact less significant (Feddema et al. 2005). In addition to climate change impacts, the LULC change is an important forcing for regional scale hydroclimatic features (Pielke et al., 2002; Pielke and Niyogi, 2010; Soloman et al., 2007). Impacts of climate change and LULC change on water resources, as a two way forcing has not been investigated to the same degree as one way forcing. In most cases, climate forcing has been changed and land-cover was kept constant (Gosian et al., 2006; Krysanova et al., 2005; Mimikou et al., 2000; Milly et al., 2005), and in some cases land-cover was changed and climate forcing was kept constant (Hurkmans et al.,2009).

Future projections of LULC changes are highly uncertain (CIMP 5). A detailed analysis of historical LULC change drivers (socio-economic versus biophysical drivers), and LULC change trajectory will be helpful for improved modeling of future LULC change scenarios. The environmental change history of the United States is relatively new (1600 to present), and well documented (Whiteny, 1994; Williams, 1989). A large geographic area of the conterminous United States (8.08 million km^2) with variable topography, soil, and climate characteristics (84 level-III ecoregions, Omernik, 1987), and the availability of high resolution cropland and population data from 1850 to present offer an opportunity to explore contributions of LULC change drivers in the United States.

1.2 Research Objectives

An interdisciplinary approach is pursued in the study by combining hydrology, atmospheric science, and land use science. The overarching goal of this study is to quantify impacts of LULC change and climate change on water resources at regional scale, and to identify major sources of uncertainty in regional water resource predictions. The historical information archived in the form

of socio-economic data (e.g. census reports) along with present day state-of-art physical modeling and analysis tools are used to develop a better/quantitative understanding of long term (past one and half century) hydrologic landscape in the United States. Three major topics pursued in this dissertation are:

(1) The Midwest Drainage Experiment: The Midwestern United States (MW-USA) was extensively drained during the second half of the 19th century and the first half of the 20th century to bring swamp and marshy land into agricultural production (The Corn Belt). The main objectives of this study are: (i) to investigate impacts of wetland drainage on the hydroclimatology of MW USA using a coupled land surface and atmospheric model, and (ii) to compare impacts of wetland drainage with impacts of climate change due to industrialization (increase in green house gas concentration). It is hypothesized that impacts of wetland drainage can be of comparable magnitude to impacts of climate change.

(2) The Water and Energy Budget Study: Studies involving both water and energy budget within a hydrologic basin (basins approach) provide insights into the closure of water and energy balance equations, and hence can lead to better conceptualization of the hydrologic system at the basin scale. The energy flux data (point scale) along with the basin approach provide the means to identify uncertainties in the reanalysis and climate model outputs. This study is focused on evaluation of North American Regional Reanalysis (NARR), and Community Land Model (CLM) outputs for surface water and energy budgets in the Mississippi River Basin (MRB). The issue of point scale observations versus climate model grid cell outputs, and model parametrization difference between NARR and CLM are also addressed in the study.

(3) The Land Cover Change History in the United States: The United States has undergone major LULC change over the past 2-3 centuries. The documented history of LULC change (1850 to present) has been analyzed along with biophysical and socio-economic drivers of LULC change. This study is focused on quantifying LULC change drivers and studying LULC change trajectory in various ecoregions of the United States. Results from this study are expected to provide a basis for designing a dynamic LULC change model for the United States as well as other parts of the world.

1.3 <u>Organization of this Dissertation</u>

This dissertation is organized as follows. Chapters 2 to 4 describe the three major topics of this dissertation in a self contained manner, i.e. each chapter has introduction, data and methods, results, discussion, and conclusion sections. A summary of findings is also presented at the end of each chapter. A synthesis study is presented in Chapter 5.

CHAPTER2. THE MIDWEST DRAINGE EXPERIMENT

2.1 Introduction

The Midwestern United States (MW USA) is comprised of eight states (Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin) covering a total land area of 1.18 million km^2 (15% of the conterminous United States). MW USA was not the land of first choice for early settlers in the region (Whitney, 1994, pp. 271-277). Presence of extensive swamp and marshes, and poor drainage network in the region (due to glacial deposits) delayed the settlement until the latter half of the 19th century (Meyer, 1935). Organizational and financial support by the government combined with technological advances spurred large scale drainage projects in the region in the latter half of the 19th century and the early half of the 20th century (Whitney, 1994, p. 274), which led to the development of the U.S. Corn Belt (comprising of Illinois, Indiana, Iowa and Ohio [McCorvie and Lant, 1993]). Approximately 20% of the total land area was artificially drained between 1870 and 1960 (Bureau of the census, 1932, 1952, 1961). Today, MW USA is an agriculturally dominated region (~60% of land cover is agricultural type), and contributes most nutrient loading to the Mississippi River, and hypoxia in the Gulf of Mexico (Goolsby and Battaglin, 2001). Large scale drainage activities carried out in the Corn Belt have resulted in loss of more than 85% of its original wetlands (Dhal, 1990).

Wetlands regulate water quality and streamflow (Novitzki, 1979), provide flood control, support biodiversity (Jenkins et al., 2003), and act as carbon sink and nutrient reservoirs. Wetlands serve as a settling area for sediment and nutrients by increasing runoff residence time, hence decreasing the nutrient load to the streams (Jones et al., 1976). Agricultural drainage of MW USA wetlands has changed the carbon balance of the region from carbon sink to carbon source (Armentano and Menges, 1986). Wetlands also play a major role in climatic feedback mechanism. Wetlands interact with the atmosphere by changing two basic land surface characteristics: surface albedo (ratio of reflected and incident solar radiations), and Bowen ratio (ratio of sensible and latent heat). Saturated soil condition (darker in color) in the upper layer in wetland allows more sun

light (or short wave radiations) to be absorbed by the land surface compared to bare soil condition (Fig. 2.1), thus reducing the surface albedo. Similarly, presence of water on the surface/ upper soil layer makes more water available for evapotranspiration (ET), thus increasing the latent heat flux, and decreasing the Bowen Ratio. A number of studies have suggested strong coupling between upper layer soil moisture and precipitation feedback [Eltahir, 1998; Findell and Eltahir, 1997; Schar et al., 1999; Koster et al., 2004; Giorgi et al., 1996] (see Section 2.2 for detail) in the present climatic condition. However, no study has been conducted to investigate hydro-climatic changes in MW USA due to large scale agricultural drainage carried out in the early part of the 20th century.

(a)   (b)
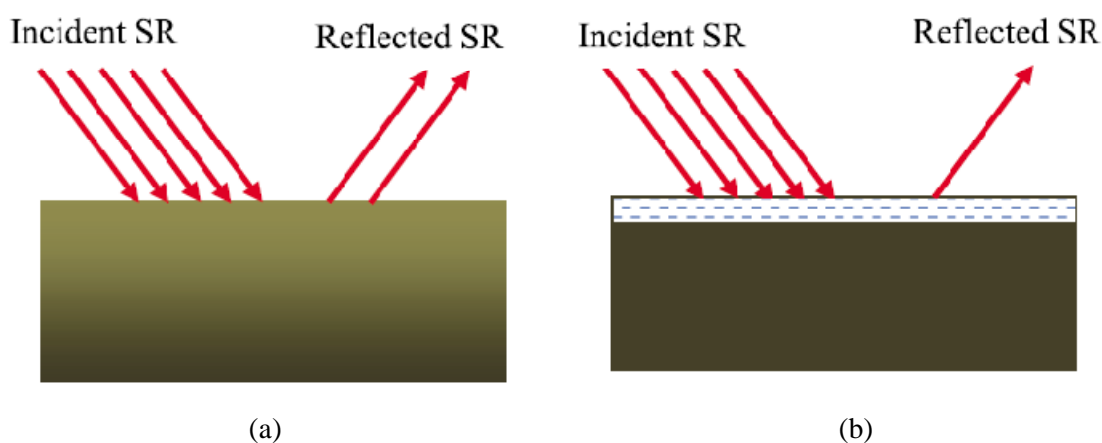
Fig. 2.1: Solar radiation (SR) interaction with: (a) bare soil, and (b) saturated upper soil layer (wetland condition)

In addition to greenhouse gas emission, LULC (Land Use and Land Cover) change is another major forcing affecting regional climate (Feddema et al., 2005; Pielke, 2005). Climatic feedback of vegetation change (forest to agricultural land) has been analyzed in the past (Bonan, 1997, 1999), but the effect of wetland drainage in MW USA has not been included in any multi-year climatic feedback simulation study. This study aims to fill this gap (investigation of the effect of wetland drainage on MW USA hydro-climatology) by using high resolution (T85 horizontal gird mesh) community climate system model to test the following two hypotheses (1) large scale artificial drainage has significantly changed the energy and water fluxes in MW USA, thus affecting the hydro-climatology (precipitation and temperature) in the region; and (2) the impact of past LULC change on MW USA hydro-climatology is comparable or even greater than the impact of impact of greenhouse gas emission based climate change in the region. These

hypotheses are tested by designing four controlled sensitivity experiments to investigate the effect of changes in CO2 concentrations and land use on MW USA hydro-climatology.

2.2 <u>Related Work</u>

Extensive literature is available on soil moisture and precipitation feedback mechanisms. While the majority of these studies suggest positive soil moisture and precipitation feedback mechanism (Eltahir, 1998; Schar et al., 1999), some studies also suggest negative soil moisture and precipitation feedback mechanism (Giorgi et al., 1996). Effect of LULC has been found to be more pronounced in summer than winter (Li and Molders, 2008). This section summarizes the findings from related work in the literature.

2.2.1   <u>Soil Moisture and Precipitation Feedback</u>

By using the observations from FIFE (First ILSCP [International Land Surface Climatology Project] Field Experiment) experiment (15 X 15 km plot size), Eltahir (1998) showed that under wet soil conditions, net radiation (short wave + long wave) at the surface increases to cause an increase in: (a) the total heat flux (sensible + latent) from the surface into the atmosphere, and (b) moist static energy supply to the planetary boundary layer (PBL), thus suggesting a positive soil moisture and precipitation feedback mechanism. In a larger scale observational study, Findell and Eltahir (1997) analyzed 14 years of soil moisture (from 19 observation sites) and rainfall data from Illinois, and found significant positive correlation between late spring/early summer soil saturation conditions and rainfall during the summer months.

Schar et al. (1999) analyzed the sensitivity of summertime European precipitation to initial soil moisture condition using a regional climate model (Europa-Model developed by German Weather Service). Their results suggested a strong coupling between initial soil moisture condition and summertime precipitation. However, majority of the surplus precipitation was not due to increased evapotranspiration in the region (Fig. 2.2a), but was caused by water vapor of advective origin that precipitated in the region due to increased precipitation efficiency under wet soil condition. Increased precipitation efficiency is due to lower rise of PBL under wet soil condition

(Figs. 2.2b and 2.2c). Findell and Eltahir (2003) used atmospheric sounding data from Central Illinois and found that there is higher probability of convection triggering (precipitation likely) under wet soil condition compared to dry soil condition.
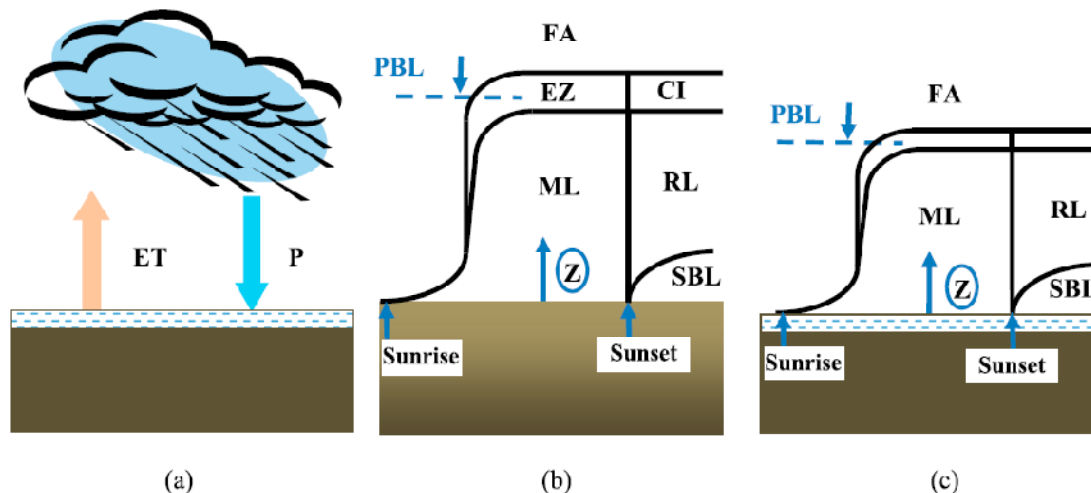
Fig. 2.2: (a) Recycling precipitation mechanism (increased evapotranspiration contributed to enhanced precipitation in the region); (b) Day time evolution of Planetary Boundary layer (PBL) in dry soil condition; and (c) Day time evolution of PBL in saturated upper horizon. ML – Mixed Layer, EZ – Entrainment Zone, FA – Free Atmosphere, SBL – Stable Boundary Layer, RL – Residual Layer, CI – Capping Inversion, Z - Height

Recycling ratio is defined as the contribution of local ET to total precipitation in the region (Fig. 2a). Using NARR (North American Regional Reanalysis) data, Dominguez and Kumar (2008) found that recycling ratio is negatively correlated to precipitation in the eastern part of MW USA (humid region), and moisture of advective origin is the largest contributor to precipitation in the region. Findings of Dominguez and Kumar (2008) are consistent with findings of Schar et al. (1999), who also concluded that increased ET under wet soil condition does not contribute directly to increased precipitation.

Giorgi et al. (1996) proposed negative soil moisture and precipitation feedback mechanism (dry condition→ increase in sensible heat flux→ increase in Buoyancy→ sustained convection) in the western part of MW USA (water limited upper Mississippi River basin). However, their findings are based on regional climate modeling experiment for two extreme cases including 1988 summer drought (La Nina event) and 1993 summer flood (El Nino event), where large scale atmospheric conditions dominate the hydro-climatology in the region. Findell and Eltahir (2003) differentiated the precipitation events dominated by large scale atmospheric conditions with

precipitation events dominated by land surface conditions in the region, and found that there are higher chances of precipitation under wet soil conditions compared to dry soil conditions in the latter case (precipitation events dominated by land surface conditions).

2.2.2   LULC change and Climate change

The term climate change in this study refers to green house gas emission based climate change, i.e., global warming due to an increase in CO_2 concentration in the atmosphere. Studying the impact of climate and LULC changes on hydro-climatology are active areas of research. LULC change is found to have significant impact on regional climate, but at global scale, increase in precipitation or temperature in one region is canceled by decreases in other regions (Feddema et al. 2005). Li and Molders (2008) used the fully coupled CCSM (version 2.0.1) to study the impact of doubled CO_2 and regional LULC change on regional and global water cycle. Impact of LULC change (very small fraction at global scale) was largely confined to the region of disturbances, and global impact of doubling CO_2 resulted in a slower water cycle (water stayed longer in the atmosphere). Numerical simulations showed that large scale (~100%) Amazonian deforestation can result in significant reduction in summer time precipitation in Amazon as well as the surrounding area (Werth and Avissar 2002). Bonan (1999) investigated the impact of deforestation in the United States using CCM3 (Community Climate Model), and found that the conversion of forest into cropland in the Eastern U.S. has decreased the mean annual surface temperature by 0.6 to 1.0 ^{O}C.

Steyaert and Knox (2008) reconstructed historical LULC change map for the Eastern U. S., and concluded that land surface characteristics have significantly changed since the pre-European settlement time (~ 1600), and suggested further investigation of climatic feedback of changed land surface characteristics. Strack et al. (2008) found that the LULC change since the European settlement in eastern U.S. has made the land surface warmer, but the precipitation is not affected for the month of June (simulations were not performed for other months). Carrington et al. (2001) investigated the impact of wetland vegetation on mid Holocene (~ 6000 years before present) climate of North Africa region using CCM3, and concluded that local recycling of water due to wetland presence was a necessary component for maintaining the mid Holocene landscape in the region covered with grassland along with scattered lakes and wetlands.

Studies reviewed in this section demonstrate the interaction between soil moisture condition and atmospheric feedback, and the impact of LULC change on regional climate. This study aims to understand these interactions due to large scale drainage activities in MW USA. In addition, unlike most previous studies on climate change that do not offer quantitative or qualitative verification of their results due to future forecasts, this study (which addresses the issue of past climate and LULC changes) uses observed climatic data over the past one century and related literature to support its conclusions. Finally, this study investigates the combined effect of climate and LULC change, as well as their individual contributions to the past hydro-climatology of MW USA. Model validation with NARR data is also expected to provide insights on the performance of CCSM3 for regional hydro-climatic studies.

2.3 Study Area

MW USA is a mid-latitude region extending from 35-50^0 N latitudes, and 80-90^0 W longitudes (Fig. 2.7). The region is mostly surrounded by the land area except for the Great Lakes in its northern part, and land surface condition plays a major role in determining regional weather and climate (Koster et al., 2000). Major river basins in the region are Upper Mississippi River Basin (UMRB), Ohio River Basin (ORB), and parts of Great Lakes Basin (GLB; Fig. 2.3a). Present landscape in the region is dominated with agricultural area covering 58% of region followed by forest (26%), wetlands (8%), and urban area (3%; Fig. 2.3b). Mean annual and seasonal climatology and their spatial variability (standard deviation) based on PRISM data (Daly et al., 1997, 1998) are presented in Table 2.1 and Fig. 2.4. Mean annual total precipitation in the region is 889 mm and annual average temperature is 8.5 ^{O}C. Mean annual temperature shows north-south gradient (Fig. 2.4a) and annual total precipitation shows north-west to south-east gradient (Fig. 2.4b). Based on climatic condition, the eastern part of MW USA (ORB; energy limited) is humid, whereas the western half (UMRB; water limited) represents the transition zone from humid to semi-arid (Berbery et al. 2003).

Glacier deposition dominates most of the landscape in MW USA (Whitney 1994, p. 44). Most recent glacial retreat in the region was 10 thousand years ago (Wisconsonian glaciation; Wikipedia 2009). Central Lowland is the major physiographic province in the region (Fenneman and Johnson 1946), which is comprised of sedimentary rock with gently sloping or undulating

topography. Glacial advance and retreat over the past two million years has muted the topography in most part of the region (Hudson 2007, p. 129). The southern and eastern boundaries of the region are un-glaciated (foothills of the Appalachians), and are characterized by plateaus and hilly terrain (Hudson 2007, page 129). Due to limited time (10 thousand years) since the last glacial retreat (on geologic time scale), most parts of the MW USA has immature topography characterized by poorly developed drainage network with presence of lakes and wetlands (Whitney 1994, page 44; Paull and Paull 1977, page 78).

(a)   (b)
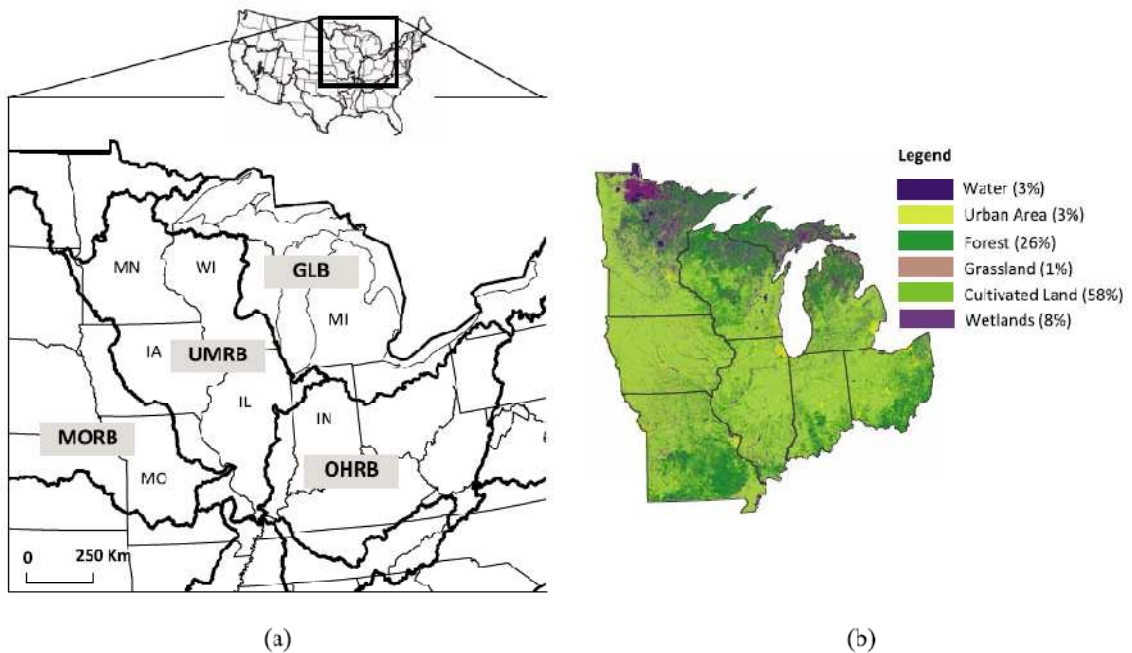
Fig. 2.3: (a) Study area (Midwestern United States) and major river basin (gray thick line); (b) Major land cover types in Midwestern United States at present. Source: NLCD 1992. GLB: Great Lakes Basin, MORB: Missouri River Basin, OHRB: Ohio River Basin, SRRB: Sourris-Red-Rainy Basin, UMRB: Upper Mississippi River Basin.

Abandonment of agricultural farms in the Eastern United States, rising demand for agricultural produce in the world market, and advancement in drainage technology (tile drains) attracted settlers to MW USA (Williams 1989, page 128; Richards, 1984). The US Congress passed the Swamp Land Acts in 1849 and 1850 for wetland reclamation. However, no large scale drainage activity could be carried out until massive state intervention came through. Passage of state drainage laws and formation of drainage districts in the latter half of the 19th century and early half of the 20th century spurred large scale drainage projects in the region (Bogue 1951; Herget 1978). Once drained, MW USA wetland soils (rich in the organic matter) became some of the

most productive soils in the world (Whitney 1994, p. 277), thus transforming areas that were once prairies and wetlands into the Corn Belt.

Table 2.1: Normal PRISM climatology for Midwestern United States.

Variable	1971-2000
Annual average T_{max} (^{O}C)	14.4(3.0)
Annual average T_{min} (^{O}C)	2.6(3.0)
Summer (JJA) average T_{max} (^{O}C)	27.5(2.1)
Winter (DJF) average T_{min} (^{O}C)	-10.5(4.5)
Annual total Precipitation (mm)	889 (151)
Summer (JJA) total Precipitation (mm)	300 (29)

Numbers represent average and standard deviation (spatial variation, in parenthesis) for the years 1971-2000. JJA: June, July and August, DJF: December, January and February.
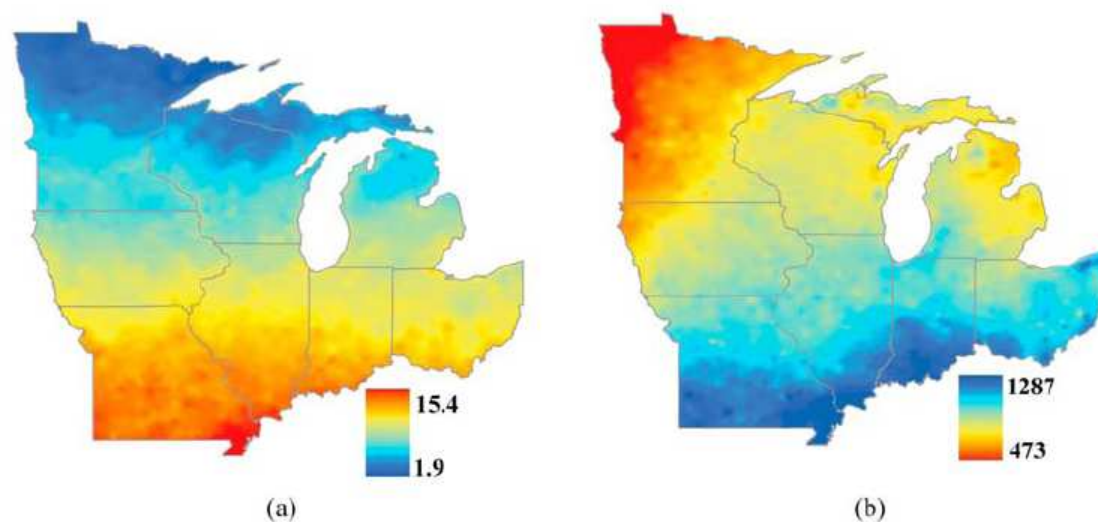
Fig. 2.4: Annual average normal climatology (1970-2000) for the study area. (a) annual average temperature in ^{O}C; (b) annual total precipitation in mm

2.4 Model Description

The Community Climate System Model Version 3.0.1 (CCSM3) is a fully coupled (atmosphere, land, ocean, and sea ice) global model to simulate earth's past, present and future climate (Collins et al. 2006). It has four independent physical components (atmosphere, land, ocean, and sea ice) which communicate with each other through a central flux coupler for exchange of fluxes

(momentum, water, heat, and salt) and state variables (temperature, salinity, velocity, pressure, humidity and air density). Internal time steps for land and atmosphere components are 10 and 20 minutes, respectively; whereas ocean and sea ice components run at hourly time step (Collins et al. 2006). The description of CCSM3 is available at http://www.ccsm.ucar.edu/models/ccsm3.0/, and its applications including model upgrades, limitations and future challenges are described in a series of papers (25 papers) in a special issue of the Journal of Climate (Gent 2006). High resolution CCSM3 (T85; ~140 km horizontal grid mesh size) has been used for future climate change projections (IPCC AR4), and low resolution (T31; ~ 375 km horizontal grid mesh size) is suitable for paleo-climate study requiring several thousand years of simulation (Yeager et al. 2006). CCSM3 and its earlier version have also been used for regional scale climate assessment/validation, climate change, and LULC change studies (Bonan, 1999; Bonan et al. 2002; Carrington et al. 2001; Dickinson et al., 2006; Li and Molders, 2008).

The atmospheric component (Community Atmosphere Model, CAM3) within CCSM3 has 26 vertical levels with pure sigma coordinates near the surface (to avoid orographic difficulty due to undulating terrain features; Phillips, 1957), hybrid sigma to pressure coordinate in intermediate zone, and pure pressure coordinates towards top of the model (above ~83mb; Collins et al. 2006). In CAM3, parameterization of clouds and precipitation processes, radiation process and aerosols are significantly revised compared to its earlier version CAM2 (Collins et al., 2006; Boville et al. 2006). The land component (Community Land Model, CLM3) within CCSM3 is integrated on the same horizontal grid mesh size as CAM3, and each gird cell has three level tile structure to represent the spatial heterogeneity of the land surface (Dickinson et al., 2006). Each gird cell is divided into four major land units: glacier, lake, wetlands, and vegetated cover including bare soil. Each land unit can have snow columns for up to five layers; each vegetation unit can have up to 4 out of 16 PFTs (Plant Functional Types), and soil columns are divided into 10 layers (Bonan et al., 2002; Dickinson et al., 2006). At every time step, CLM outputs include surface albedo, upward long- wave radiation, sensible and latent heat fluxes, and surface wind stress to CAM. In turn CAM provides incident solar radiation, incident long-wave radiation, convective and large-scale precipitation, lowest level temperature, horizontal wind components, specific humidity, and pressure height above surface to CLM at each time step (Zeng et al. 2002). Wetlands are parameterized as a standing column of water (no soil) without vegetation. No explicit surface runoff, infiltration and sub-surface drainage are allowed from the wetland, and runoff from wetlands is calculated as a residual of the water balance term (Oleson et al., 2004).

2.5 <u>Methodology</u>

The methodology involves conducting four sensitivity experiments to investigate the effect of climate and LULC changes in MW USA using CCSM3. The effect of climate change is incorporated by using different CO_2 concentration levels, and the change in LULC is incorporated by using historical drainage data. Long term historical observations for temperature and precipitation, and related work in the literature have been used to verify conclusions drawn from the model results. Key points in the methodology are elaborated below.

2.5.1   <u>Drainage Data Reconstruction</u>

Spatial and temporal distribution of wetland drainage estimates are required for the CCSM3 sensitivity experiments. Drainage estimates from the United States Department of Commerce Bureau of Census reports (Bureau of the Census, 1932, 1952, 1961, 1973, 1981) are used in this study. Bureau of the Census provides decadal time series of state level drainage estimate from year 1870 and county level drainage estimate from year 1920. Decadal time series of total drainage area (Land in Drainage Enterprises) for Midwest and conterminous USA is shown in Fig. 2.5. As evident from Fig. 2.5b, MW USA remains the most drained region, with approximately 20% of its area drained by year 1960 compared to approximately 5% for the conterminous USA during the same time period. In 1970s, new large scale drainage project activities slowed considerably because of various reasons including (i) most farms requiring drainage through large scale public project had been fully or partially drained, (ii) technological advancements placed private drainage activity within the reach of individual farmers, and (iii) change in governmental policy (Bureau of the Census, 1978). From 1977 onwards, the federal government stopped supporting new drainage projects, thus ending the era of government supported large scale drainage projects in the United States (McCorvie and Lant, 1993). Although some wetland restoration activity has begun, it is much less (0.1%) compared to massive scale drainage activity carried out earlier (McCorvie and Lant 1993). Based on the present day land use data (NLCD 1992), it is reasonable to assume that the massive scale of drainage projects installed in early half of 20th century has been maintained on the landscape, as all the drained area remains in agriculture production in the Midwest region (Figs. 2.3b and 2.6a).
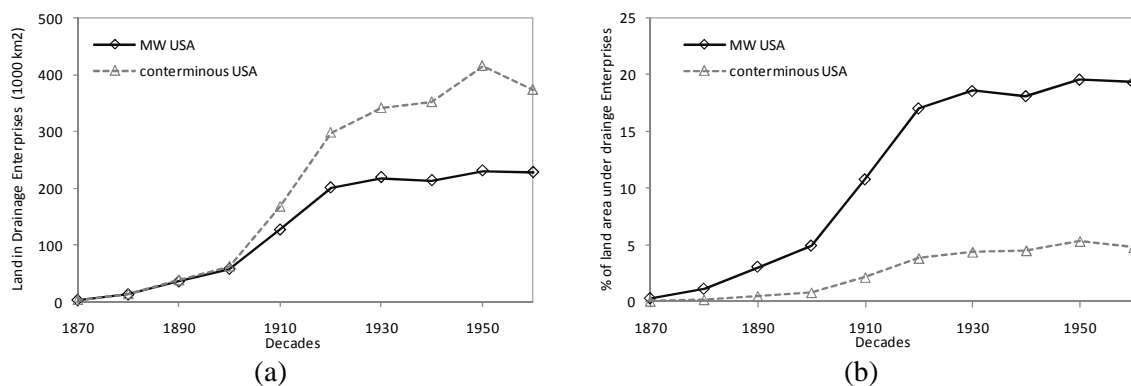
Fig. 2.5: Decadal time series of drainage extent in United States and MW USA region (a) Absolute land area (b) percentage of total land area

Fig. 2.6a shows the spatial variability in drainage estimate (expressed as % of area drained) for year 1930. After 1930, drainage area in the Midwest reached a plateau; whereas the drainage area at the national level continued to increase at a slower pace (see Fig. 2.5). Midwest region exhibits a high degree of spatial variability in drainage intensity (Fig. 6b): Indiana is the most highly drained state (~48% in year 1960) followed by Ohio (~33%), Michigan (~27%), Minnesota (20%), Iowa (19%), Illinois (15%), Missouri (~7%) and Wisconsin (~2%). Spatial variability in drainage areas in different states could be due to availability of land area requiring drainage in individual states. For sensitivity experiment design, the drainage estimate for 1950 is used because it represents the maximum drainage area in MW USA (Fig. 2.5). The county level drainage intensity map is first rasterized to T85 grid cell size (~ 140 km), and overlapped with CCSM3 grid to assign wetland percentage to individual grid cells in the Midwest region. Fig. 2.7a shows the present (1990's) wetland extent in CCSM3 surface data (default data set), and Fig. 2.7b shows the reconstructed wetland data for 1870 condition. Average wetland representation in the masked domain (shown in Fig. 2.7) for the present condition is 1%, and for 1870 condition, it is 19%.

Information related to wetland loss is also available from Dhal (1990), Pavelis (1987), and Steyaert and Knox (2008). Dhal (1990) estimated 14% wetland loss in the Midwestern states from 1780 and 1980. One of the possible reasons of difference between Bureau of the Census estimate (~20%) and Dhal (1990) estimate (14%) could be the definition of wetlands. A farm may require drainage to improve soil-environment in the root zone, and it may not come under wetland definition (Pavelis 1987, Dideriksen et al. 1978). Potential saturated soil (~ wetlands) maps reconstructed by Steyaert and Knox (2008) are consistent with Dhal (1990). Because

Bureau of the Census provided the most detailed estimate (spatial distribution at county level) compared to other two datasets (Dhal (1990) at state level and Pavelis (1987) at national level), Bureau of the Census estimate is used for sensitivity experiment design.
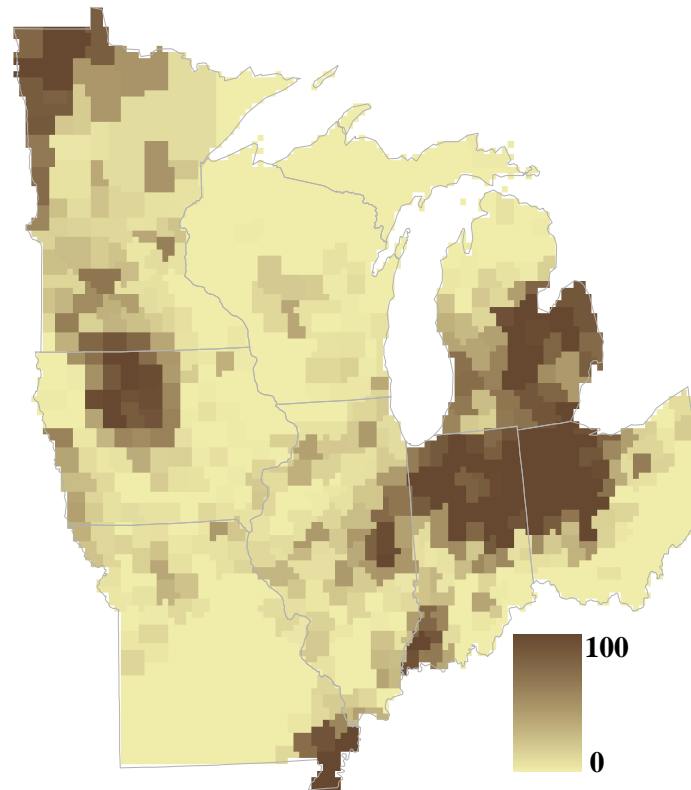
Fig. 2.6 (a): Drainage intensity map for year 1930's. (Raster map of % of county area drained)
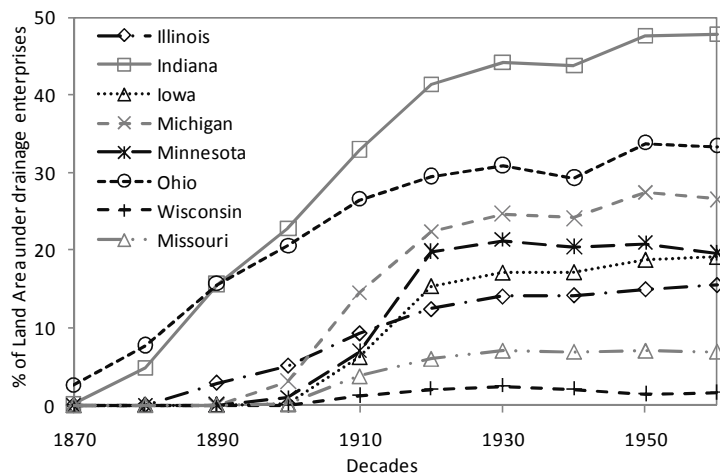
Fig. 2.6 (b): Decadal time series of drainage extent in Midwestern States (% of land area drained)

2.5.2   Sensitivity Experiment Design and Model Runs

In this study interaction between atmosphere (CAM) and land surface (CLM) is modeled at T85 (~140 km) horizontal gird mesh size with default sea surface temperature (SST) and sea ice concentration at 1° horizontal gird mesh size (Hurrell SST climatology). Because the primary focus of this study is to analyze the impact of wetland drainage (LULC change) at regional scale (MW USA only), default SST and sea ice concentration are used. This approach is consistent with past studies by Bonan (1997; 1999) and Findell et al. (2009). In addition, SST is found to have stronger influence in tropic and sub tropic regions than in mid latitude where land-atmospheric interaction plays a greater role (Koster et al. 2000). High resolution CCSM3 (T85 horizontal grid mesh) is used in this regional scale assessment study because it has been found to show the highest climate sensitivity (increase in global average mean surface temperature when atmospheric CO2 concentration is doubled; Collins et al., 2006) compared to intermediate (T42 - horizontal grid mesh) and low resolution (T31 horizontal grid mesh) CCSM results (Kiehl et al., 2006; Kiehl and Gent 2004).

To compare the effects of wetland drainage with green house gas based climate change effects for MW USA (regional scale), the model is run for 1990 (355 ppm) and 1870 (289 ppm) CO2 concentration levels with and without wetlands in each case. Except for wetland change in surface data (Fig. 2.7), and green house gas based external forcing (CO2 concentration), all other datasets (e.g., vegetation type, SST) are kept unchanged across different modeling experiments.

Thus, a total of four control runs are designed for the study (i) 355 ppm CO2 for year 1990 excluding wetland (1990C), (ii) 355 ppm CO2 for 1990 including wetland (1990W), (iii) 289 ppm CO2 for year 1870 excluding wetland (1870C), and (iv) 289 ppm CO2 including wetland (1870W). For each experiment, the model is run for 25 years including the first five years as spin-up period. Thus, only 20 years of monthly model output from each experiment is included in the analysis.
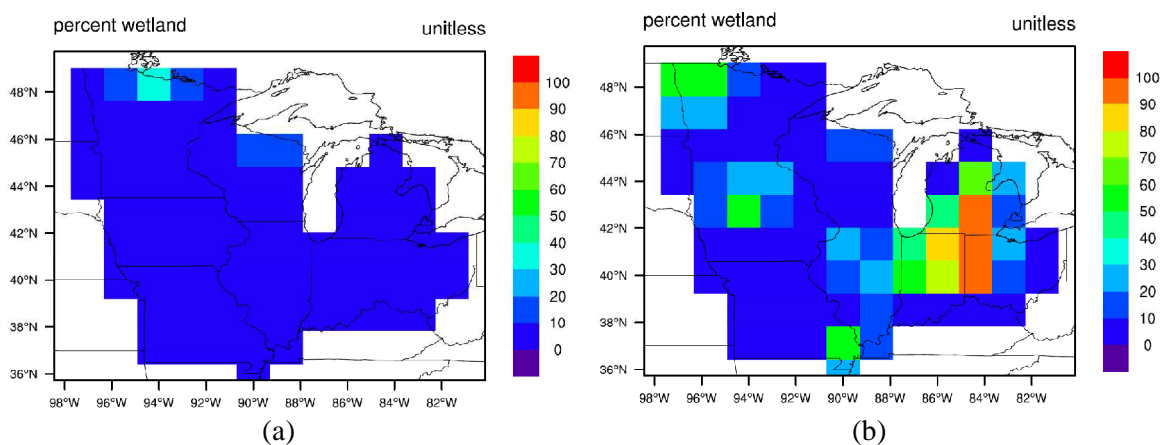
Fig. 2.7: Wetland representation in CCSM3 (a) present condition in year 1990 (1%) (b) Pre-settlement condition in year 1870 (19%)

CCSM3 model runs are executed on the Purdue TeraGrid Steele Cluster through a newly developed user friendly web portal (http://www.purdue.teragrid.org/ccsmportal). Any user can take advantage of TeraGrid computing resource for running CCSM3 model through this portal (Basumalik et al., 2007). Model output is processed using NCL (NCAR Command Language, http://www.ncl.ucar.edu/) and NCO (netCDF Operators, http://nco.sourceforge.net/).

2.5.3   Model Validation

Before analyzing the results of sensitivity experiments, the CCSM3 model results for existing conditions (1990C) are validated against high resolution NARR (North American Regional Reanalysis) data. NARR provides a high frequency (3-hour), high resolution (32 km mesh size) atmospheric and land surface hydrology data set for the North American domain from 1979 to present. The NARR dataset is developed as a major improvement upon earlier National Center for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) Global

Reanalysis (GR) datasets in terms of resolution and accuracy (Mesinger et al. 2006). NARR data are generated by using lateral boundary conditions from NCEP GR2 (Kanamitsu et al., 2002), the NCEP Eta model (Mesinger 2000) and its data assimilation system (Rogers et al., 2001), the NOAH land surface model, and observations (e.g. precipitation, radiance, 10 m wind seed). Successful assimilation of high quality detailed precipitation observation in the NARR system provides a much improved dataset of land-atmospheric interaction and land-surface hydrology (Mesinger et al. 2006).

The NARR data are re-gridded to T85 mesh size, and monthly NARR climatology for the present time is prepared from a 20 year monthly data set (1980-1999) centered around 1990. Analysis is performed for the MW domain only by applying a regional mask to the re-gridded NARR data. Fig. 2.8 shows the monthly climatology of the NARR data versus CCSM3 results (1990C) for selected variables. Annual average values and temporal variability from NARR data and all four CCSM3 runs are presented in Table 2.2. Summer is considered from May to October (6 months), and winter is considered from November to April (6 months).

Monthly climatology of incoming solar radiation (FSDS) at the surface is well simulated by CCSM3. There is a negative bias in winter (-24.7 watts/m^2) and small positive bias in summer (3.7 watts/m^2), thus causing an annual under estimation of 5% (-10.5 watts/m^2). Net short-wave radiation is also well simulated by CCSM3, but a small overestimation during the summer creates an overall annual overestimation of 7% (10.7 watts/m^2). There is a large overestimation of net long-wave radiation flux during the summer months to produce an annual overestimation of 27% (16.3 watts/m^2).

Sensible heat flux from CCSM3 is highly overestimated (32% or 8.7 w/m^2), and latent heat flux is highly under estimated (32% or -20.9 w/m^2) during the summer. NARR climatology of sensible heat flux and latent heat flux is similar to the climatology found by Berbery et al. (2003) using NCEP's Eta model. Berbery et al. concluded that during the summer (coinciding with the vegetation growing season), increases in net radiation (short- and long-wave) are balanced by increases in latent heat flux with minimal or no increase in sensible heat flux (Figs.2.8b-c). However, CCSM3 results do not show considerably enhanced (~ 2 times) latent heat flux compared to sensible heat flux during the summer. Improper partitioning of net radiation into sensible and latent heat flux may be one of the main reasons behind other summer biases

observed in CCSM3 results (e.g. warming bias during the summer, overestimation of net long-wave radiation flux).

Table 2.2: Mean annual climatology of NARR and CCSM3 results for MW USA.

Variable	Unit	NARR	1990C	1990W	1870C	1870W
Incoming solar radiation	watts/m^2	210.2 (13.0)	190.7 (13.6)	189.5 (14.4)	193.1 (14.6)	192.7 (14.5)
Net short wave radiation	watts/m2	146.4 (11.2)	157.1 (11.7)	157.4 (12.4)	158.8 (12.4)	159.2 (12.6)
Net long wave radiation	watts/m^2	60.6 (8.4)	76.9 (10.1)	72.4 (9.6)	77.8 (10.7)	73.1 (9.5)
Sensible heat flux	watts/m^2	27.0 (9.4)	35.8 (12.4)	26.1 (10.8)	36.5 (12.6)	27.7 (10.7)
Latent heat flux	watts/m^2	64.8 (9.8)	43.9 (12.8)	58.4 (10.9)	44.1 (13.1)	57.9 (10.7)
2 m air Temperature	^{O}C	9.5 (2.1)	11.7 (2.2)	11.2 (2.2)	11.2 (2.4)	10.3 (2.2)
Total Precipitations	mm/year	839 (431)	615 (356)	667 (363)	616 (371)	656 (352)
Convective Precipitation	mm/year	420 (235)	286 (176)	334 (173)	288 (176)	333 (170)
Large Scale Precipitation	mm/year	419 (278)	329 (180)	333 (190)	328 (195)	323 (182)
ET	mm/year		551	773	553	727
P-ET	mm/year		64	-106	63	-71

Numbers in parenthesis represents average values of monthly standard deviation. For precipitation monthly standard deviation is multiplied by 12 to be consistent with annual total value.

(a) Incoming solar radiation at surface

(b) Net short wave radiation flux

(c) Net long wave radiation flux

(d) Sensible heat flux

(e) Latent heat flux

(f) 2 m air temperature

(g) Total precipitation

(h) Difference of summer precipitation
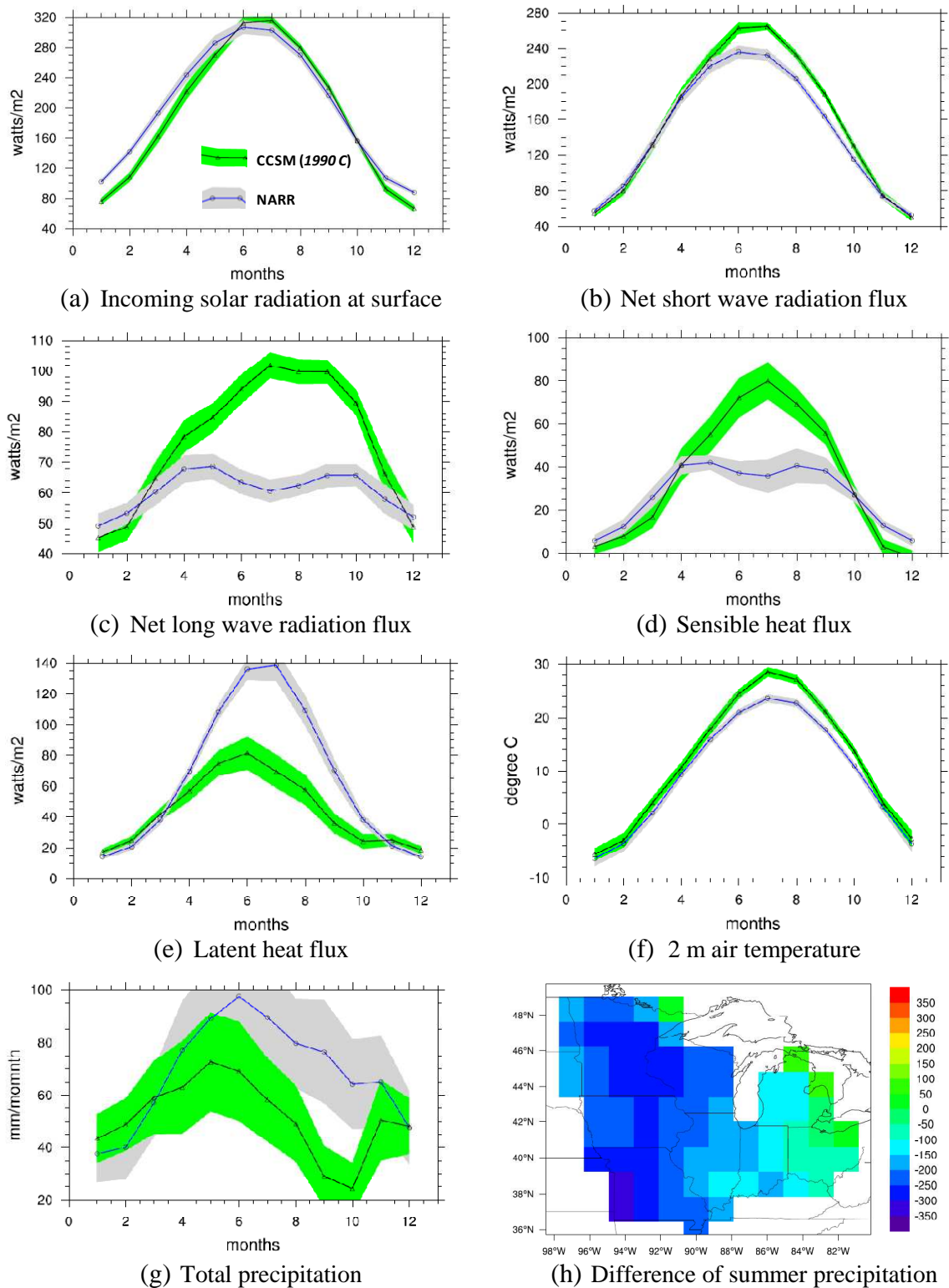
Fig. 2.8: Comparison between NARR climatology and CCSM3 climatology for 1990C case. Shaded regions show inter annual variability at 95% confidence level in each case

The monthly climatology of 2 m air temperature is well captured by CCSM3. The 2 m air temperature is overestimated during summer months to produce an annual warming bias of 2.3 ^{O}C (Fig. 2.8f). It should be noted that observed 2m air temperature is not assimilated in the NARR system. However, NARR produced improved prediction of 2m air temperature compared to GR2 data (Mesinger et al. 2006). One of the major success stories of NARR is assimilation of observed precipitation; hence NARR precipitation can be treated as observed data (Mesinger et al. 2006). Compared to NARR data, CCSM3 highly underestimates summer precipitation to produce an annual underestimation of 25% (-206mm) (Fig. 2.8g). A closer look at precipitation data shows that convective precipitation dominates the summer precipitation. Therefore, convective precipitation is more underestimated (32% or -132 mm) compared to large scale precipitation (22% or -91 mm). Total precipitation is equally divided between convective and large scale precipitation in NARR as well as in CCSM3 outputs (Table 2.2). Similar findings of CSSM results compared with the observed precipitation data are also reported for central and eastern US by Bonan et al (2002), and for the St. Lawrence region (area surrounding Great Lakes in USA and Canada) by Li et al. (2008).

The spatial distribution of summer precipitation difference shows that underestimation of precipitation is confined mainly to the western region of the study domain (Upper Mississippi and Missouri river basin); whereas overestimation (although small) is confined to the eastern domain (Ohio River basin). Underestimation of large scale precipitation (-80 mm MW summer total) is evenly distributed across the MW domain, but convective precipitation (-108 mm MW summer total) is highly underestimated in the western domain with small overestimation of convective precipitation in the eastern domain (not shown). It seems that the CCSM3 convective scheme works fine in the energy limited (abundance of soil moisture for ET) eastern domain, but it highly underestimates precipitation in the water limited (availability of water becomes a limiting factor for evapotranspiration) western domain (Berbery et al., 2003; Dominguez and kumar, 2008). Li et al. (2008) reports the shortcoming of convective scheme as one of the possible reasons for precipitation underestimation in CCSM3 results.

2.5.4   Examination of Inter-annual Variability in Model Fluxes

Validation of certain CCSM3 output variables (e.g., net long wave radiation, sensible and latent heat flux) show that CCSM3 results have significant biases. It can be argued that this model bias can be handled in CCSM3 sensitivity experiments by taking the difference of model outputs. To support this argument, and to gain more faith in the difference between two CCSM3 outputs, CCSM3 results are examined for their ability to describe inter-annual variability in model energy fluxes between dry and wet years. The mean difference in monthly climatology between the five driest and five wettest years from 1980 to 1999 using NARR data and CCSM3 results for the study region is presented in Table 2.3 and Fig. 2.9. CCSM3 climatology is extracted from the seven member ensemble CCSM3 20th century run for the study region (available at Earth System Grid, http://www.earthsystemgrid.org/). Table 2.3 presents both the annual difference in climatology and the difference during summer months (May to October). Seasonal data are presented because previous studies (e.g., Koster et al., 2000; Li and Molders, 2008) have found that land surface condition in the mid-latitude region has considerable effect on hydro-climatic variables during summer months.

Table 2.3: Mean difference climatology between 5 driest and 5 wettest years during 1980 to 1999 for NARR and CCSM3 data in MW USA.

Variable	Units	Annual		May – October	
		NARR	CCSM3	NARR	CCSM3
Net short wave radiation	watts/m^2	4.2	7.0	5.1	5.8
Net long wave radiation	watts/m^2	4.3	7.3	6.0	7.3
Sensible heat flux	watts/m^2	6.8	8.1	11.0	10.9
Latent heat flux	watts/m^2	-5.2	-8.5	-10.7	-13.0
2 m air Temperature	^{O}C	0.20	0.79	0.30	1.57
Total precipitation	mm	-201	-215	-118	-119

CCSM3 values represents mean of seven member ensemble runs. Summer Average is for the month May to October.

(a) Net short wave radiation

(b) Net long wave radiation

(c) Sensible heat flux

(d) Latent heat flux

(e) 2m air temperature

(f) Total precipitation
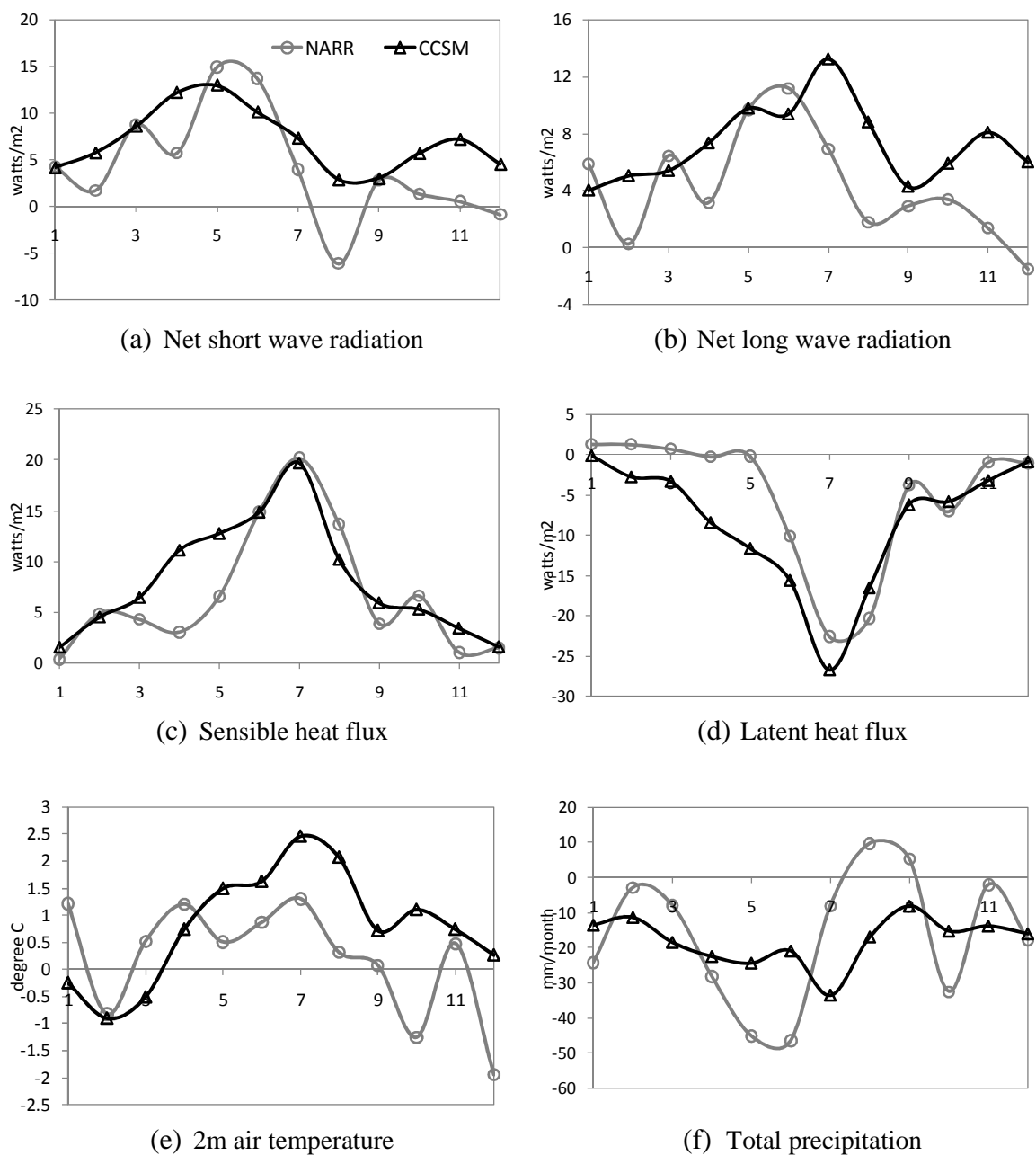
Fig. 2.9: Difference climatology of NARR and CCSM3 between 5 driest and 5 wettest years during 1980 to 1999 for the study region. CCSM3 climatology represents mean value of 7 member ensemble run (Section 2.5.4). X-axis represent month January to December (1 to 12).

The results show that difference in total annual precipitation between the driest and wettest years is almost equal for both NARR (-201mm) and CCSM3 (-215mm) as shown in Table 2.3. The overall pattern as well as monthly variations in most climatic variables is comparable between NARR data and CCSM3 ensemble runs as shown in Fig. 2.9. The magnitude of changes in surface energy and radiation fluxes are of same order for both NARR data and CCSM3 output. The annual difference in surface energy and radiation flux is higher in CCSM3 results compared to NARR data (63% for latent heat flux). However, results are much better for the summer months compared to the annual data (21% for latent heat flux during summer). These differences could be due to difference in land surface parametrization scheme and/or may be some biases in NARR data could also be contributing to the diffrences. CCSM3 shows higher sensitivity to temperature change compared to NARR data between dry and wet years with both annual and seasonal data. The overall similarity of CCSM3 outputs in capturing the inter-annual variability in most climatic variables, particularly during summer seasons, strengthen the argument that the results obtained by taking the difference in CCSM3 model output can provide reasonable findings by canceling out the model biases.

2.6 Sensitivity Experiment Results

2.6.1   LULC change versus climate change

The results of the effect of climate and land use change are presented by subtracting the CCSM3 output for one scenario from another to neutralize the effect of biases found in CCSM3 during validation. It is argued that the biases observed during validation will apply equally to all CCSM3 results, and hence the difference between two outputs will be unbiased (Li and Molders, 2008; also discussed in Section 2.5.4). In addition, for the most part climatic variability is well captured by CCSM3 results (width of shaded region in Figs. 2.8 a–g and standard deviations shown in Table 2.2); hence, statistical significance of the difference in climatology can be ascertained. Global evolution of climate under changed land cover conditions is expected to provide better assessment (dynamically consistent) in CCSM3 experiments compared to a RCM (Regional Climate Model) based study where lateral boundary condition are taken from either

observation/reanalysis or GCM (Global Climate Model) output that are not forced with changed landscape condition (e.g. Strack et al., 2008).

Results from CCSM3 runs are combined in 5 different groups, with each group representing the difference between two CCSM3 runs. These five groups are: (1) 1990C-1990W: effect of wetland drainage under 355 ppm CO2 concentrations, (2) 1990W-1870W: effect of climate change with CO2 concentration increase from 289 to 355 ppm in the presence of wetlands, (3) 1990C-1870W: combined effect of wetland drainage and climate change, (4) 1870C-1870W: effect of wetland drainage under 289 ppm CO2 concentrations, and (5) 1990C-1870C: effect of climate change with CO2 concentration increase from 289 to 355 ppm without the presence of wetlands. For most variables, results from Groups 1 and 4, and Groups 2 and 5 are found similar, and therefore results from only groups 1-3 are presented in detail. Figs. 2.10 -2.16 show the results from CCSM3 runs for the effects of wetland drainage (1990C-1990W), climate change (1990W-1870W), and combined climate change and wetland drainage (1990C-1870W) in terms of difference in monthly climatology for the Midwest domain. Annual average values and their temporal variability (standard deviations) calculated from 20 years of monthly output data for each CCSM3 run are presented in Table 2.2.

Net short-wave radiation is the difference between incoming solar radiation and reflected solar radiation at the surface (+ve downward). As demonstrated in Fig. 2.1, drainage of wetlands for agricultural activities will increase the reflected solar radiation and decrease the net short wave radiation, thus increasing the surface albedo. Decrease in net short wave radiation due to wetland drainage is predominant during most of the summer months (May to September); however, it is found statistically significant only during the month of June (Fig. 2.10a). The effect of climate change is the decrease in net short wave radiation over most part of the year; however, no result is statistically significant (Fig. 2.10b). Annual average decrease in net short wave radiation due to wetland drainage is 0.3 watts/m^2 (+ve and –ve sign during winter and summer months cancel each other), and decrease due to climate change is 1.8 watts/m^2. Thus, the combined effect of climate change and wetland drainage on net short wave radiation is a reduction of 2.2 watts/m^2. Decrease in net short wave radiations due to combined effect is statistically significant during July and August (Fig. 2.10c). The spatial distribution of decrease in net short wave radiation due to combined effect during the summer months (Fig. 2.10d) follows the wetland drainage pattern (Fig. 2.7b), with higher decrease in wetland drained cells compared to small or no change in non-wetland drained cells.
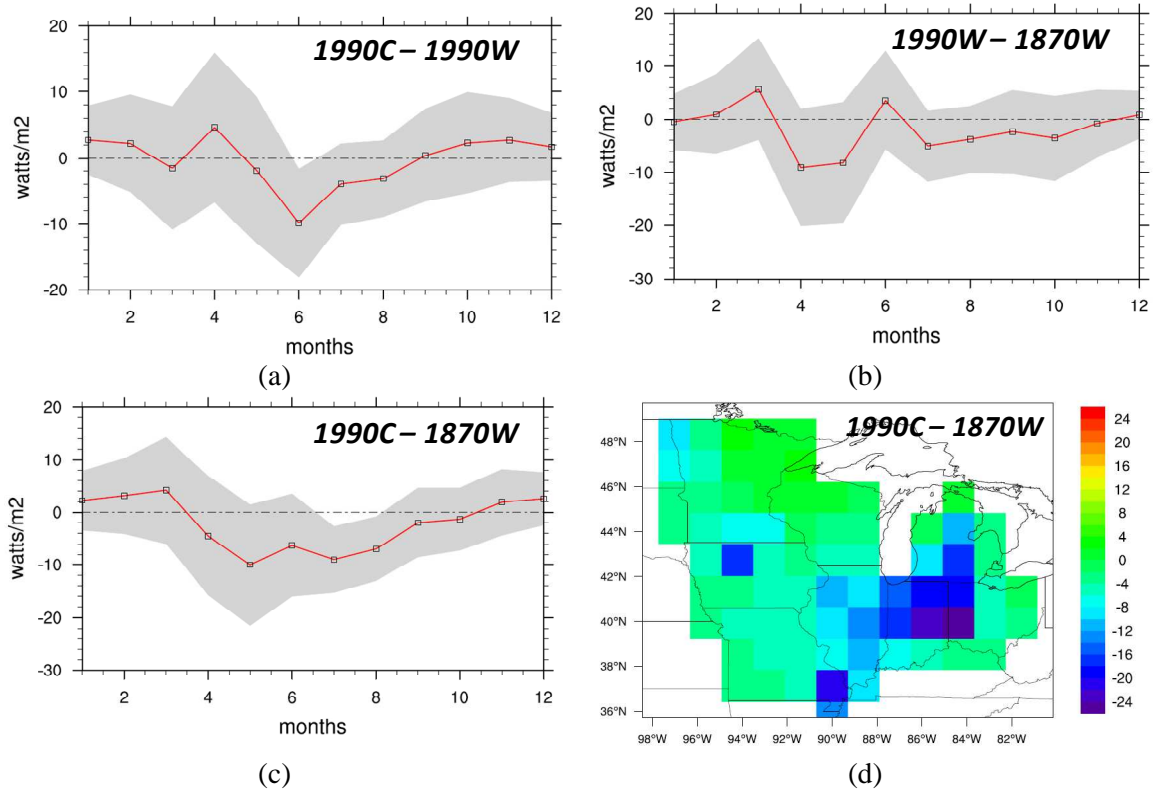
Fig. 2.10: Difference in Net short wave radiation (watts/m²) between CCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

Net long wave radiation is the difference between the outgoing long wave radiation emitted by the earth surface and the atmosphere (terrestrial radiation) and the downward long wave radiation from the atmosphere to the earth surface (+ve upward). Warmer earth surface due to wetland drainage has increased net long wave radiation, particularly during the summer months, and the results are statistically significant from July to October (Fig. 2.11a). The effect of climate change is a small decrease in net long wave radiations during most part of the year, but the results are not statistically significant for any month (Fig. 2.11b). Annual average increase in net long wave radiations due to wetland drainage is 4.4 watts/m²; whereas annual average decrease in net long wave radiation due to climate change is 0.7 watts/m². Therefore, the combined effect is an increase of 3.7 watts/m² increase in annual net long wave radiations. Monthly climatology of the combined effect follows the monthly climatology of the effect of wetland drainage, and statistically significant increases are found during the months of July to September (Fig. 2.11c). The spatial distribution of increase in net long-wave radiation during the summer months (Fig.

2.11d) follows the wetland drainage pattern (Fig. 2.7b), with higher increase in wetland drained cells compared to small or no change in non-wetland drained cells.
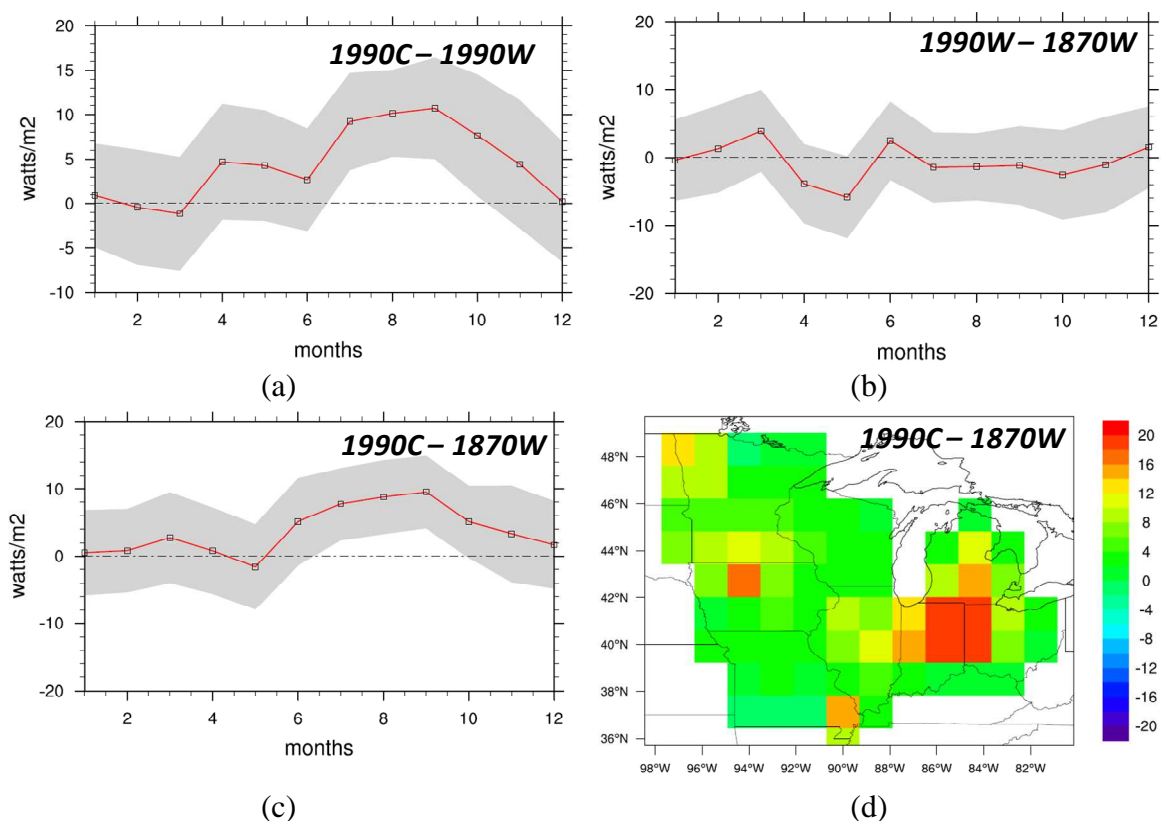
Fig. 2.11: Difference in Net long wave radiation (watts/m^2) between CCCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

Sensible heat flux has increased during most part of the year due to wetland drainage, and the results are statistically significant from April to October (Fig. 2.12a). There is no significant change in sensible heat flux due to climate change (Fig. 2.12b), and the combined effect follows the pattern similar to that of the wetland drainage (Fig. 2.12c). Annual average increase in sensible heat flux due to wetland drainage is 9.7 watts/m^2; whereas the annual average decrease in sensible heat flux due to climate change is 1.6 watts/m^2. Therefore, the combined effect is an as annual average increase of 8.1 watts/m^2 in sensible heat flux. The spatial distribution of the increase in sensible heat flux due to combined effect during the summer months is shown in Fig. 2.12d, which is similar to the distribution pattern from wetland drainage as shown in Fig. 2.7b.
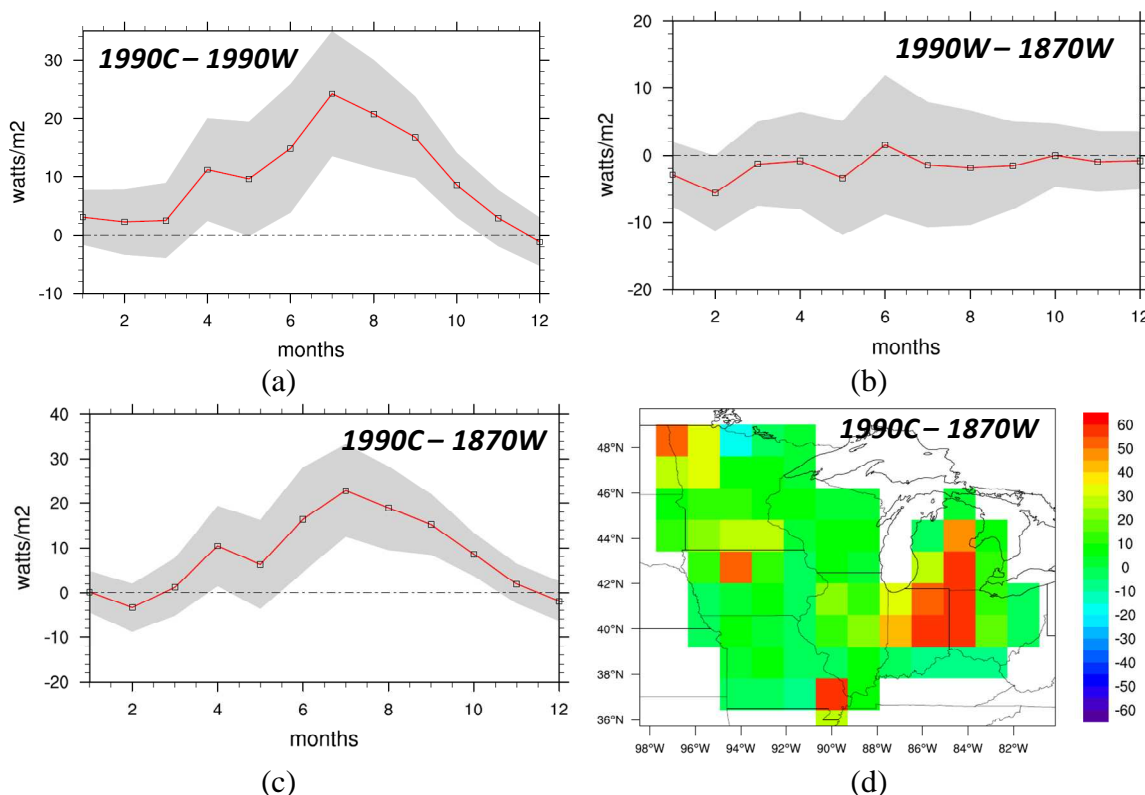
Fig. 2.12: Difference in sensible heat flux (watts/m^2) between CCCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

Latent heat flux has decreased during most of the year due to wetland drainage, and the results are statistically significant for April- October (Fig. 2.13a). There is no significant change in latent heat flux due to climate change (Fig. 2.13b), and the combined effect follows the pattern similar to that of wetland drainage (Fig. 2.13c). Annual average decrease in latent heat flux due to wetland drainage is 14.5 watts/m^2; whereas the annual increase in latent heat flux due to climate change is 0.5 watts/m^2. Therefore, the combined effect is an annual average decrease of 14.0 watts/m^2 in latent heat flux. Magnitude of annual decrease in latent heat flux is higher (14.0 watts/m^2) compared to annual increase in sensible heat flux (8.1 watts/m^2), thus compensating for decrease in net radiation (short wave (-2.2) + long wave (-3.7) = -5.9 w/m^2). The spatial distribution of the decrease in latent heat flux due to the combined effect is shown in Fig. 2.13d, and it is similar to the distribution pattern from wetland drainage (Fig. 2.7b). Ground heat flux constitutes a minor component of surface energy flux, and considering long term equilibrium

condition, ground heat flux (which is not presented in this study) can be taken as zero (Berbery et al., 2003; Dirmeyer et al., 2006).
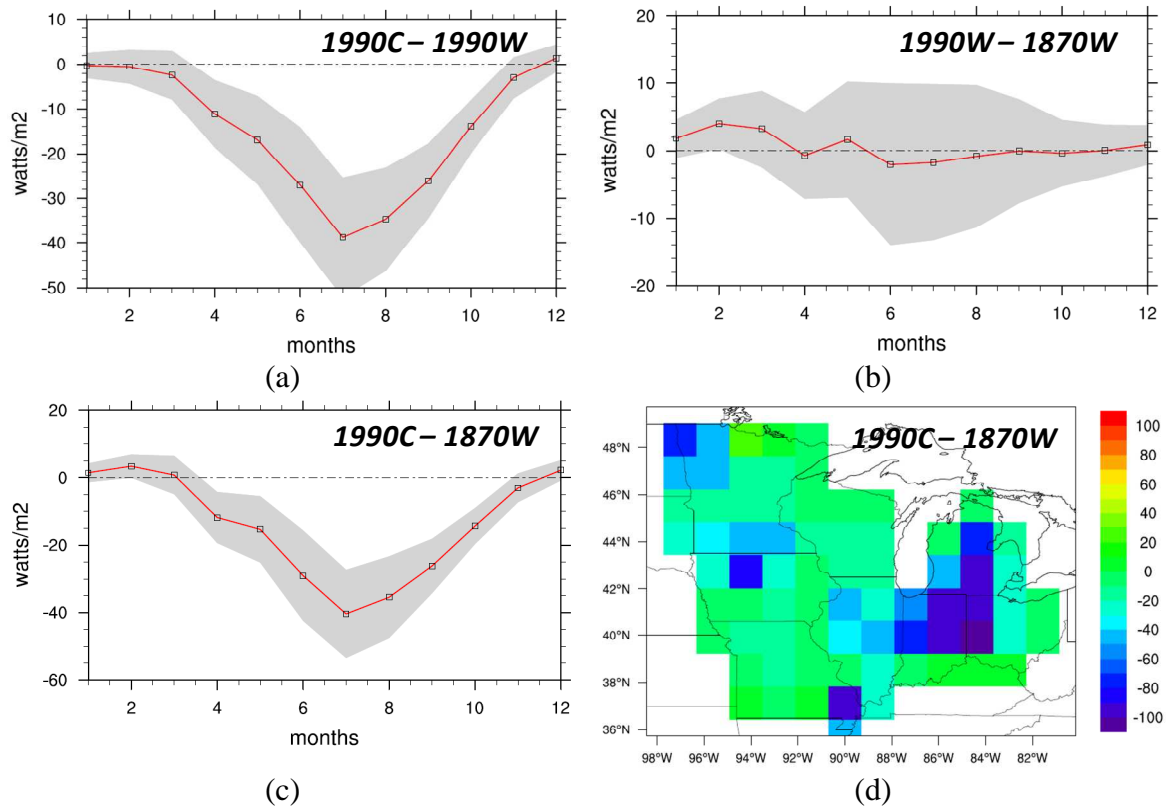
Fig. 2.13: Difference in latent heat flux (watts/m^2) between CCCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

Two meter air temperature has increased during summer months due to wetland drainage, and the results are statistically significant for July and August (Fig. 2.14a). Climate change has resulted in increased winter temperature that is statistically significant from January to March (Fig. 2.14b). As a result, the combined effect is a statistically significant increase in 2m air temperature during most of the year (Fig. 2.14c). Annual average increase in 2m air temperature due to wetland drainage and climate change is 0.6 ^{O}C and 0.9 ^{O}C, respectively. As a result, the combined effect from wetland drainage and climate change is 1.4 ^{O}C. The increase in 2m air temperature for wetland drained cells is relatively higher compared to un-drained cells in all summer (Figs. 2.14d and 2.17a) and winter (November-April; not shown) months.
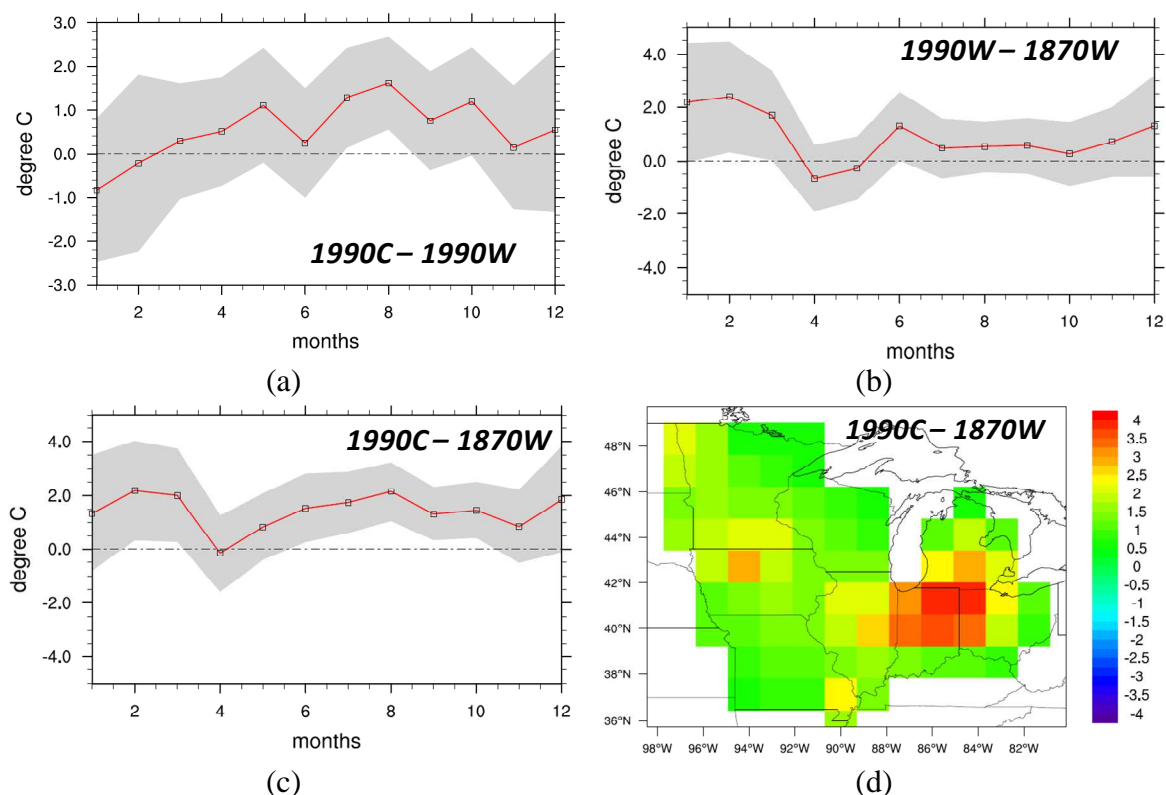
Fig. 2.14: Difference in 2m air temperature (^{O}C) between CCCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

The 1870C-1870W scenario that represents the effect of wetland drainage under 289 ppm CO_2 concentration results in an increase in annual temperature (0.9 ^{O}C); whereas the 1990C-1870C scenario that represents the increase in CO_2 concentration from 289 ppm to 355 ppm under no wetland condition results in an annual increase in temperature (0.5 ^{O}C), which is not statistically significant for any month (not shown). Average increase in annual temperature due to wetland drainage (0.8 ^{O}C, average of 1990C-1990W, and 1870C-1870W) is equivalent to average increase in annual temperature due to climate change (0.7 ^{O}C, average of 1990W-1870W, and 1990C-1870C).

Convective and large scale precipitation contributes equally to the total precipitation in the region (Table 2.2). Convective precipitation dominates during the summer season and large scale precipitation dominates during the winter season (not shown). Convective precipitation has decreased due to wetland drainage during the summer months, but it is statistically significant only for the month of July (Fig. 2.15a). Climate change has not resulted in any significant change

in convective precipitation (Fig. 2.15b), and the combined effect follows the climatology of wetland drainage (Fig. 2.15c). Annual total decrease in convective precipitation due to wetland drainage is 52 mm, and annual total increase in convective precipitation due to climate change is 11 mm. Therefore, the total annual decrease in convective precipitation due to combined wetland drainage and climate change is 41 mm. The change in the magnitude of convective precipitation is proportional to the increase in the percentage of drainage in grid cells (Fig. 2.17b). For grid cells representing 50% or more drainage extent, the decrease in summer convective precipitation is more than 100 mm (Fig. 2.17b). Under 1870 CO_2 concentration (1870C-1870W), the effect of wetland drainage on convective precipitation is similar to that of 1990 condition (1990C-1990W), but the results are not statistically significant for any month (not shown). There is no statistically significant change in large scale precipitation in the region either due to wetland drainage, or due to climate change (Fig. 2.16).

(a) (b) (c) (d)
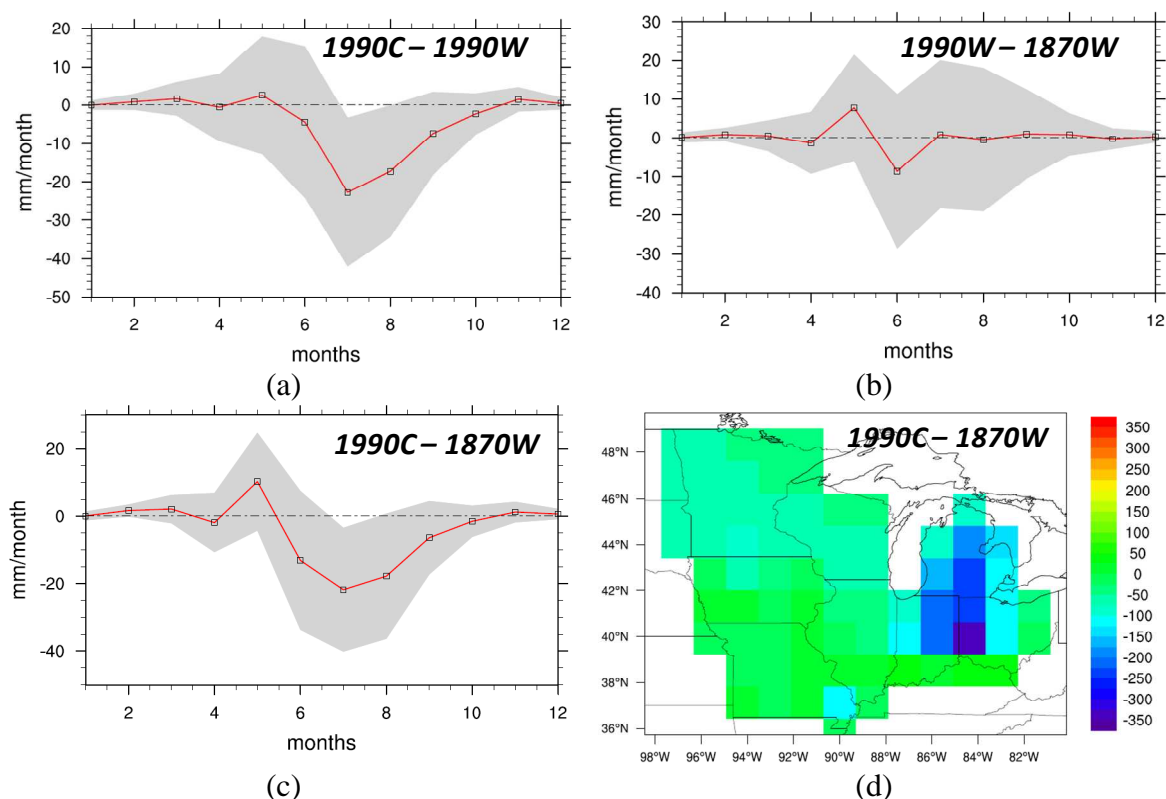
Fig. 2.15: Difference in convective precipitation (mm) between CCCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

(a)   (b)

     (c)   (d)
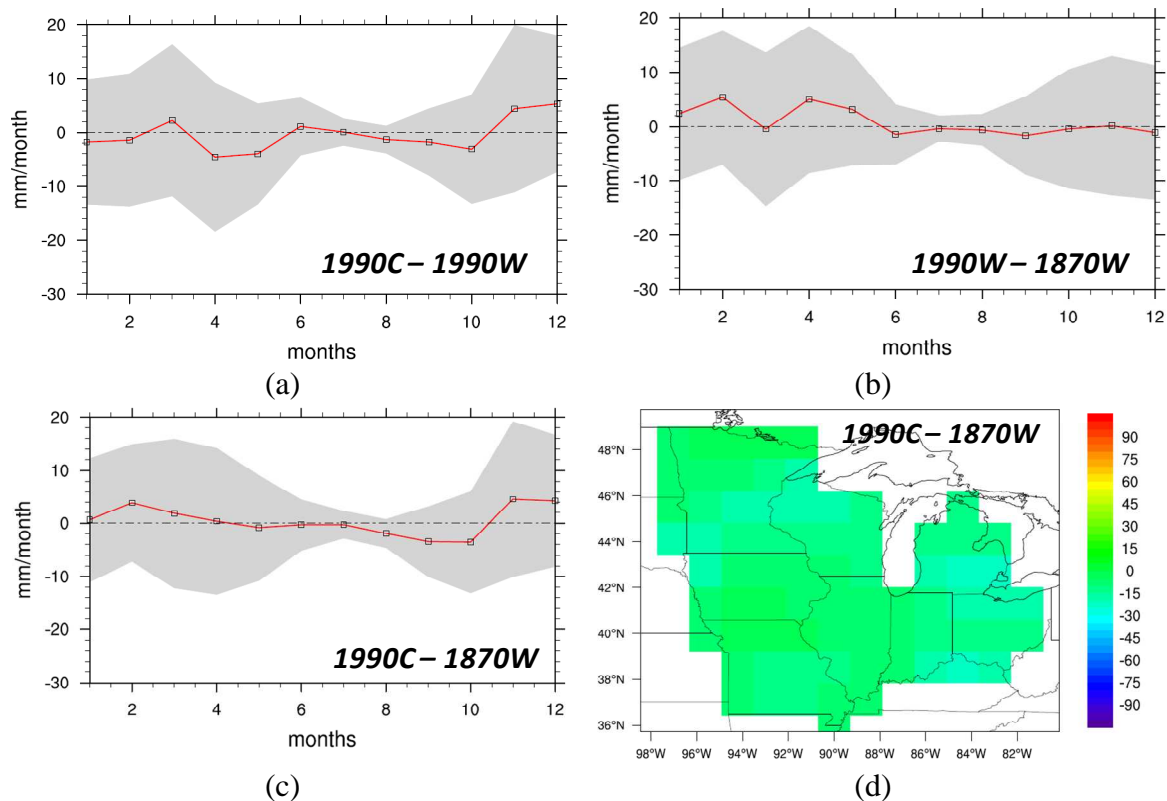
Fig. 2.16: Difference in large scale precipitation (mm) between CCCSM3 runs. Shaded region in (a) – (c) represents 95% confidence level uncertainty range calculated from 20 years of monthly data from each CCSM3 run. Spatial distribution for each variable in (d) represents absolute difference over the summer months (May to October).

(a)   (b)
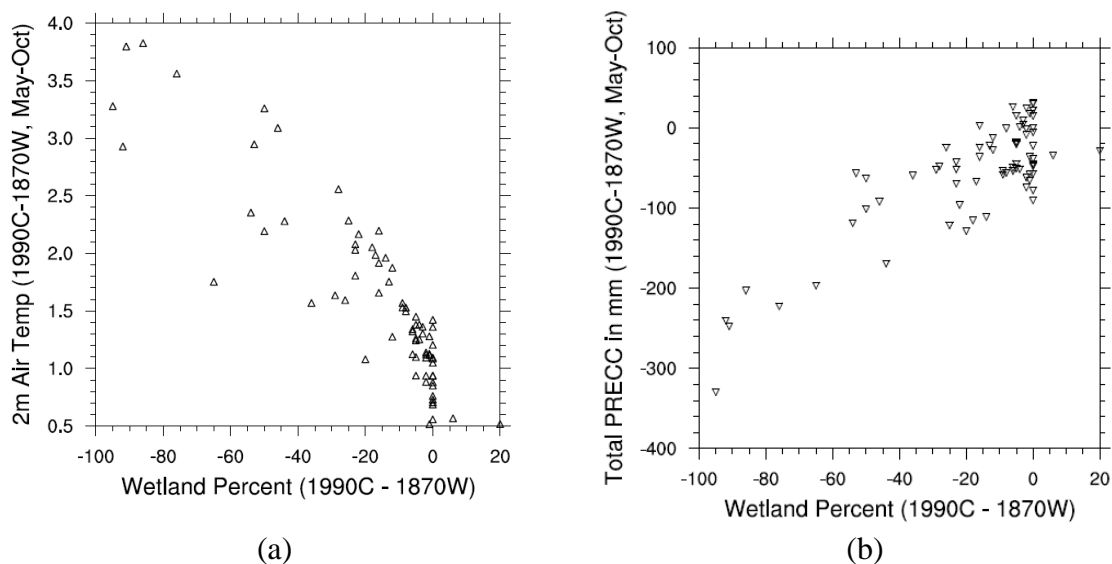
Fig. 2.17: Scatter plot of the combined effect (1990C-1870W) on (a) Summer 2 m Air Temperature, (b) Summer Total Convective Precipitation.

2.6.2   Long term water balance (Precipitation – Evapotranspiration)

Long term (20 years in this study) annual average of the difference between precipitation and evapotranspiration (P-ET) represents blue water availability in the region (Gordon et al., 2003; Schuol et al. 2008). Blue water is defined as the water that is available for consumptive purposes (i.e., the sum of surface and sub-surface runoff and deep aquifer percolation, Schuol et al. 2008). Wetland drainage has increased P-ET mainly due to the decrease in ET. Climate change has resulted in a relatively small decrease in P-ET, mainly contributed by increases in ET over precipitation. Annual total increase in P-ET due to wetland drainage is 170 mm, and the total decrease in P-ET due to climate change is 35 mm. Therefore, the combined effect is an increase in blue water availability (by 135 mm) in the region (Table 2.2). Under 1870 CO2 concentration (1870C-1870W), the effect of wetland drainage on P-ET is similar with slightly smaller magnitude (134 mm). Under no wetland condition (1990C-1870C), climate change has not affected the overall water availability.

P-ET serves to define long term maintenance of wetlands in the region (Carrington et al., 2001). P-ET is -106 and -71 mm for 1990W and 1870W, respectively, and 64 and 63 mm for 1990C and 1870C, respectively (Table 2.2). Spatial distribution of annual P-ET for 1870W and 1990C scenarios is shown in Fig. 2.18. Negative values of P-ET for wetland cells indicate that water from adjoining cells is required for maintenance of wetlands in the region. Partitioning of total ET into ground evaporation (E_g), canopy evaporation (E_c), and transpiration from plants (E_{tr}) is shown in Table 2.4. Ground evaporation constitutes a major portion (73%) of total ET under wetland conditions (1870W), and it reduces to 62% under non-wetland conditions (1990C). ET exceeds precipitation in both cases (153 mm in 1870W and 56 mm in 1990C) during summer months.
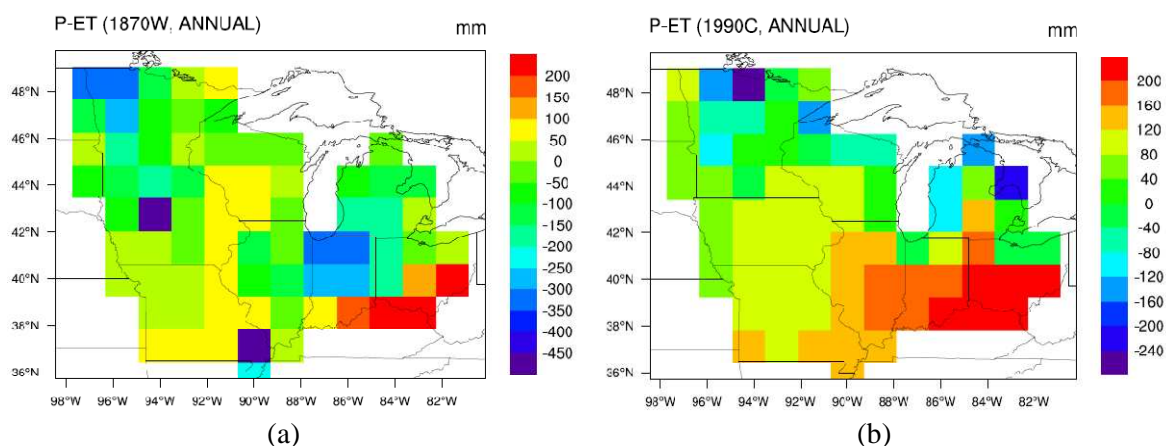
Fig. 2.18: Spatial distribution of annual P-ET (a) for 1870W case (b) for 1990C case.

Table 2.4: Partitioning of Evapotranspiration into its components: Ground evaporation (E_g), Canopy evaporation (E_c), and Transpiration (E_{tr}), (Annual/Summer total in mm).

Variable	Annual		Summer	
	1990C	**1870W**	**1990C**	**1870W**
E_g	340	530	182	360
E_c	133	127	117	112
E_{tr}	78	70	71	63
ET (Total)	551	727	370	535
P	615	656	314	382
P-ET	64	-71	-56	-153

2.7 <u>Discussion and Implications</u>

CCSM3 model was used to investigate the impact of historical wetland drainage on Midwest hydro-climatology. The effect of wetland drainage is compared with the effect of green house gas emission based climate change between year 1870 and 1990. Modeling results suggest that wetland drainage has resulted in significant changes in regional energy budget (sensible and latent heat flux) and radiation budget (long wave radiation), particularly from May to October. As a result, the climate has become warmer, and convective precipitation has decreased during summer months. Except for 2 m air temperature, climate change has not resulted in any significant changes in other important climatic variables (e.g. precipitation, sensible heat flux, latent heat flux, net short and long wave radiation) in the region. Results from this study highlight

the importance of LULC change compared to climate change in MW USA over the past one and half century.

Validation of CCSM3 results with NARR data for the Midwest domain points to some of the shortcomings in CCSM3. The partitioning of net radiation into sensible and latent heat flux during summer months is imprecise in CCSM3 for the study region. CCSM3 largely overestimates sensible heat flux, and underestimates latent heat flux during summer months. One possible reason could be that CLM3 may not be capturing the crop dynamics (increased ET during summer/growing season; Chow et al., 1988, p. 91-93) properly. Another reason could be that the underestimation of summer precipitation is getting reflected in underestimation of latent heat flux (Lawrence et al., 2007). However, Wang et al. (2008) also found similar results (underestimation of latent heat flux) based on offline simulation (forced with observed climatic inputs) of CLM3 for two different land cover sites (rain forest and agricultural site). Wang et al. (2008) found that the soils in CLM3 are excessively dry (due to higher runoff) and show much less seasonal variation as compared to observed data. The issue of partitioning of net radiation into sensible and latent heat flux should be investigated in more detail because it has further implications in terms of warming bias in the summer, overestimation of net long wave radiation, and lower precipitation efficiency (Schar et al., 1999).

Higher underestimation of summer precipitation (> 200 mm) in CCSM3 results is found to be concentrated in the western domain of the study region. When CCSM3 results are explored for the entire United States, it is found that there is a narrow strip (between 90^0 West and 100^0 West longitude) of highly underestimated summer precipitation that lies between the drier western part and the wetter east part (not shown). In semi-arid to arid western part and humid eastern part of the United States, precipitation simulation are much better (within $\pm$ 100 mm). The transitional zone between wet and dry climatic region has been identified as a hot spot for studying the influence of soil moisture on precipitation (Koster et al., 2004). One possible reason could be the interaction between the excessively drier soil in CLM3 (Wang et al. 2008) and its feedback to precipitation (underestimation) in the transitional zone (Koster et al., 2004), but this issue needs further investigation.

Previous studies have found biases in global reanalysis data (Berg et al., 2003; Maurer et al., 2001; Roads and Betts, 2000), and it is reasonable to suspect some biases in NARR data as well. Therefore, some of the biases found in CCSM3 results in this study could also be due to the

NARR data, and difference in land surface parameterization scheme. Investigation of biases in NARR data is the scope of another study, presented in Chapter 3. Despite the biases in CCSM3 output, a comparative analysis of CCSM3 flux estimates from driest and wettest years from 1980 to 1999 show that these model biases cancel out when one output is subtracted from the other, thus providing confidence in the results of sensitivity experiments conducted in this study.

For long term maintenance of wetlands in the region, annual P-ET should be close to zero (Carrington et al., 2001). The negative P-ET (e.g. -71mm for 1870W case) observed in this study could be the result of a combination of factors such as: (1) boundary of the study domain, (2) wetland and ET parameterization in CCSM3, and (3) underestimation of summer precipitation. The Midwest domain selected in this study is based on political state boundaries (simplified by creating Midwest mask, see Fig. 2.18) instead of using hydrologic or watershed boundary (shown in Fig. 2.3a) for the region. When great lakes and other surrounding areas were included in the analysis (full area shown in Fig. 2.18, instead of only Midwest region), difference between annual precipitation and ET was slightly positive (18mm; not shown). In CCSM3, wetlands are parameterized as open water surface without vegetation, and this results in higher ground evaporation (Table 2.4). However, inclusion of vegetation in wetland would have decreased total ET, and simultaneously, it would have also decreased precipitation to balance the difference between precipitation and ET (Carrington et al., 2001). Issues related to ET parameterization (excessive ground evaporation and canopy interception) are discussed by Lawrence et al. (2007). Use of improved version of land model (CLM3.5; Oleson et al. 2008) might have resulted in better simulation result (positive P-ET). Higher underestimation of precipitation (>200 mm) particularly in the summer could have also contributed to negative value of P-ET in the region.

Sensible and latent heat flux are found to be the first order variables affected by the drainage of wetlands in the Midwest region. Sensitivity experiments in this study showed significant increase in sensible heat flux, and decrease in latent heat flux during summer months. Although the biases in CCSM3 may have played some role in these results, the annual average change in fluxes simulated in this study is several times greater than the uncertainty indentified in CCSM3 results (Figure 2.9 and Table 2.3). For example, annual average decrease in latent heat flux due to wetland drainage is 14.5 watts/m^2; whereas uncertainty in CCSM3 results in simulating the difference in latent heat flux between dry and wet years is 3.3 watts/m^2 (Table 2.3). Hence, it is reasonable to conclude that changes in land surface condition have significantly affected surface energy fluxes in the region. Magnitude of changes may differ (can become smaller) depending

upon the model parameterization of wetlands (e.g. Carrington et al., 2001), but it seems less likely that either the sign or their statistical significance will be affected. Sensible and latent heat fluxes are the major energy source to the atmospheric boundary layer, and are responsible for modulating weather conditions in the region (Bonan, 2008; p. 217-222). The importance of sensible and latent heat flux for regulating local/regional weather condition makes these variables critical in studying the impact of land use change at local to regional scale.

Hydro-climatological impact was not confined to highly drained cells only (say > 40%; Fig. 2.19), thus suggesting a regional scale impact of the wetland drainage. Fig. 2.19 shows Students' t values (at 95% confidence level) for July 2 m air temperature and convective precipitation. Statistically significant impact is distributed across all the wetland drainage at all levels (0-100%). However, the percentage of wetland drainage correlated well with the changes in temperature and precipitation. Average monthly correlation coefficients for temperature and precipitation change (1990C-1870W) during summer months (May to Oct.) are –0.81 and 0.64, respectively (not shown). During winter months (November to April), 2 m air temperature was slightly less correlated with wetland drainage percentage (average monthly correlation coefficient: -0.63; not shown). The effect of wetland drainage was largely confined to the Midwest USA region, suggesting no major teleconnection effects (not shown). This result is consistent with Findell et al. (2009) who also found that the effect of land use change was confined to the region of disturbance.

CCSM3 results suggest an annual warming of 1.4 $^\text{O}$C due to the combined effect of wetland drainage and climate change. Summer warming is mainly caused by wetland drainage; whereas winter warming is caused by green house based climate change. However observed data show that over the last 112 years (1896-2006), Midwestern climate has become warmer by 0.44 $^\text{O}$C (annual average temperature change) mainly due to increase in winter temperature (1.08 $^\text{O}$C during Dec-Feb, and 0.65 $^\text{O}$C during Mar-May). The change in summer temperature is negligible or slightly negative (0.03 $^\text{O}$C during Jun-Aug, and -0.15 $^\text{O}$C during Sept-Nov; http://mcc.sws.uiuc.edu/climate_midwest/mwclimate_change.htm). The higher increase in winter temperature is consistent with CCSM3 output in this study (significant increase in winter temperature due to climate change). However, the significant increase in summer temperature from CCSM3 output is not consistent with observed data. It should be noted that vegetation type was kept unchanged from the 1990C condition [dominantly cropland] for all modeling experiments, and wetland insertion caused an 18% decrease in vegetation cover (replaced by

wetlands). Bonan (1997, 1999) has studied the role of vegetation change and found that conversion of forest to cropland has cooled the midwestern climate during summer (0.5 ^{O}C during Jun-Aug, and 2.5 ^{O}C during Sept-Nov; Bonan 2001). It seems that observation of summer temperature change is somewhere in between the warming due to wetland drainage and cooling due to vegetation change. Inclusion of vegetation change along with wetland drainage and better parameterization of wetlands (wetlands with vegetation; Carrington et al., 2001) might result into more realistic summer temperature simulation.
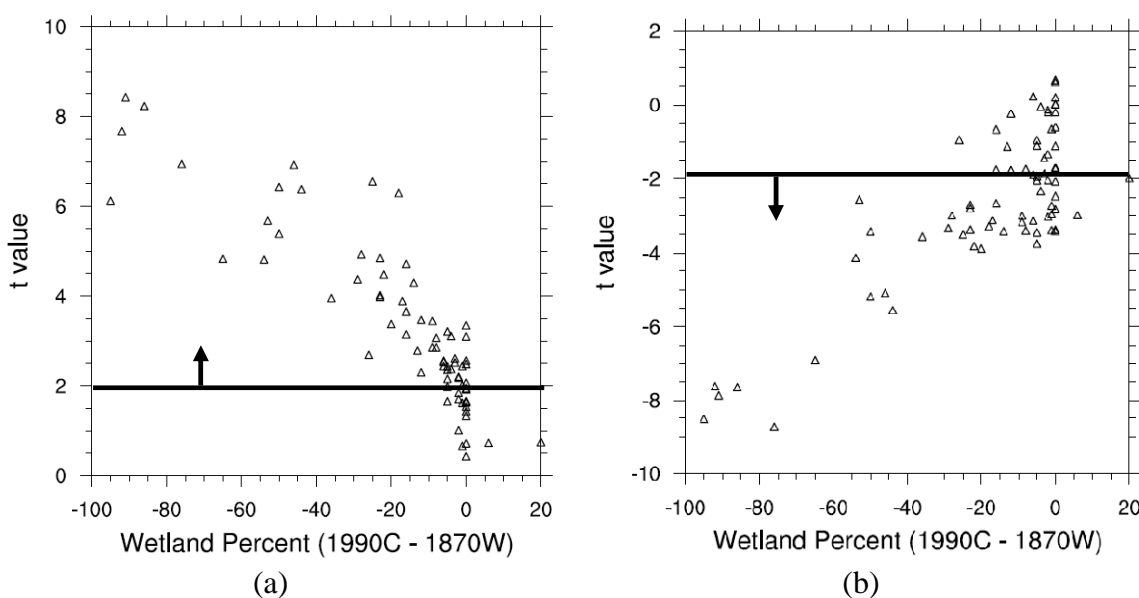
Fig. 2.19: Scatter plot of statistical significance (t value) of difference in July month (a) 2 m air temperature, and (b) convective precipitation for 1990C-1870W case (combined effect). Results above and below black think line in (a) and (b) respectively are statistically significant (marked by arrow sign).

CCSM3 results suggest an overall decrease of 41 mm in annual precipitation due to the combined effect of wetland drainage and climate change. This decrease is mainly contributed by reduced precipitation in summer months due to wetland drainage. However, observed data show an annual increase of 74 mm in Midwest precipitation from 1896-2006 (http://mcc.sws.uiuc.edu/climate_midwest/mwclimate_change.htm). The annual precipitation increase is distributed among three seasons (Mar-May: 20 mm, Jun-Aug: 24 mm, Sept-Nov: 27 mm) with almost no increase during the winter (Dec-Feb: 2mm). However, the data length (1896-2006) is not sufficient to see the effect of wetland drainage on precipitation, as wetland drainage was already underway by 1900, and also precipitation over the last century would have been affected by large scale climatic changes.

Based on some early observations (Schott, 1881), Wahl (1968) has compared the climate of the early half of the 19th century (1830's and 1840's) with climate from 1931-1960. Wahl found that the Midwest region was wetter (20-30% higher precipitation during the summer and early fall) during the early half of the 19th century. However, Wahl attributed wetter climatic condition to large scale climatic features (moisture influx from the gulf, and polar cold fronts). It may be a good time to revisit the Wahl findings (particularly attribution factor) in view of new knowledge about the effect of wetland drainage on precipitation. Fig. 2.20 shows the mean annual total precipitation before year 1876 observed at stations having 18 years and longer data records (average: 26 years, and maximum: 48 years; Schott, 1881) along with the current normal precipitation (1970-2000) data from PRISM (Daly et al., 1997, 1998). The available data from Schott (1881) captures the north-west to south-east precipitation gradient in the region. Overall data from Schott (1881) does not show any significant change in precipitation pattern in the region. However, no long term observed data (before year 1876) are available in highly drained pocket (intersection of IN, OH and MI states) to see the effect of wetland drainage on precipitation.
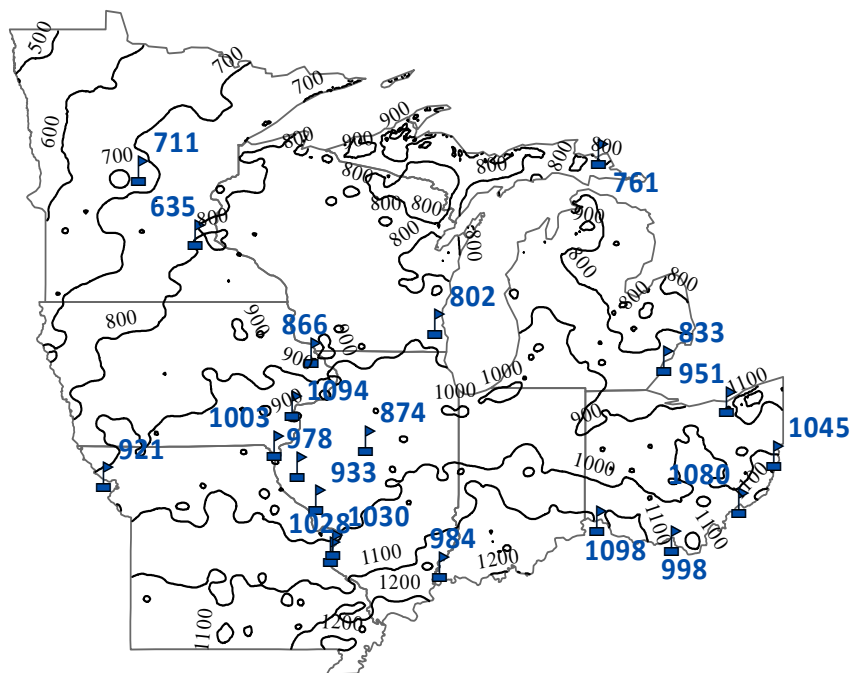
Fig. 2.20: PRISM normal annual precipitation (in mm, 1971-2000) along with Smithsonian Institute precipitation data for 19th century (number in bold blue letters; Schott, 1881) of average data length 26 years (minimum: 18 and maximum 48 years) before 1876.

The extensive network of sub-surface tile drains in the Midwest USA quickly drains the water infiltrating in to the upper soil layers (Kumar et al. 2009; Basu et al. 2009), and prevents the rise of water table to the crop root zone. The water table in MW USA is relatively shallow (Fan et al., 2007), and contributes towards evapotranspiration demand (via capillary rise) during summer months (Miguez-Macho et al., 2007; Yeh and Famiglietti 2009), which may suggest that the landscape is not substantially different, and the wetland drainage has not really caused significant hydro-climatic changes of the region. However, in the wetland condition upper soil layers are saturated during most part of the year and vegetation is adapted for saturated soil conditions (USEPA; http://www.epa.gov/owow/wetlands/what/definitions.html), which is not the case with an agricultural landscape. Hence, the present agricultural landscape in MW USA is different from the pre-settlement landscape. Saturation level in upper soil layers significantly affects land surface characteristics (Eltahir, 1998), suggesting significant hydro-climatic changes in MW USA due to wetland drainage.

To draw an analogy of magnitude of changes simulated in this study with similar changes observed in the recent past; changes in latent heat flux and 2m air temperature between a drought year (1988 with total annual rainfall of 670 mm) and flood year (1993 with total annual rainfall of 954 mm) from NARR data are shown in Fig. 2.21, along with the simulated effect of wetland drainage (1990C – 1990W). Soil saturation for upper soil layers was quite high (0.70 -0.80) in year 1993, and it was very low (0.30 to 0.40) in year 1988 particularly during the summer months (soil moisture observation for top 30 cm soil layer in Illinois are provided by [Findell and Eltahir 1997]). The magnitude of change between year 1988 and 1993 is comparable or even greater than the magnitude of change due to wetland drainage (1990C-1990W) found in this study (Fig. 2.21). Hence, it is reasonable to assume simulated changes in this study are not unrealistic.

Effects of wetland drainage on surface energy fluxes, precipitation, and temperature found in this study are consistent with similar experiments using regional climate models for shorter duration ranging from few days to months (e.g., Nebraska irrigation simulation study by Adegoke et al., 2007; effect of land use change by Marshall et al., 2004). By using a coupled Regional Atmospheric Modeling System and a plant model for central grassland region (South Dakota, Nebraska, Kansas, Wyoming, and Colorado) in the United States, Estaman et al. (2001) also found that the effect of land use change was more pronounced at regional scale compared to effects from increased CO2 concentration (2 x CO_2 concentration). The results found in the present study can be used as a basis for more detailed investigations (diurnal cycle of temperature

and precipitation events) using a fine resolution regional climate model. Regional climate models are reported to do a better job in representing land surface heterogeneity, and simulating weather events because of their different model physics (e.g. convective scheme) and finer resolution (Han and Roads, 2004; Liang et al., 2004 and 2006). By using a global climate model in a multiyear simulation study, the present study is able to convey a larger message that at regional scale, the impact of land use change is comparable or even greater than the impact of green house gas emission based climate change.

(a) Latent heat flux   (b) 2m air temperature
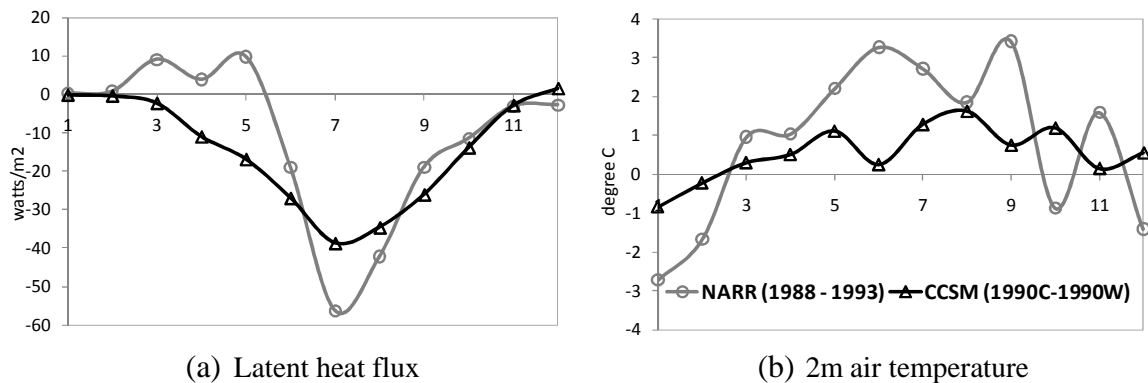
Fig. 2.21: Difference between two contrasting land surface condition 1988 (dry) and 1993 (wet) for NARR data along with the effect of wetland drainage simulated in this study for MW USA region. X-axis represents number of month 1 to 12 (January to December).

In climate and LULC change studies, analysis of the difference climatology between control runs and the run designed to study the impact of change is a standard practice (e.g., Bonan 1997, 1999; Eastman et al., 2001; Findell et al, 2009; Li and Molders 2008). This study also follows the same approach and argues that taking the difference between two control runs cancels out the modeling biases for most climatic variables, Nevertheless, the issue of model biases and its effect on the overall findings needs further work including detailed investigation and validation of individual modeling components of climate models at regional scale. For example, the drier summer in CCSM3 output (Fig. 2.8g) could lead to higher surface temperature, higher outgoing long wave radiations, and higher sensible heat at the expense of lower latent heat and lower precipitation efficiency. Similarly, CLM3 (Land Component of CCSM3) may not be capturing the crop dynamics properly. Addressing issues like these in climate modeling, and their effect on hydroclimatic predictions needs more attention for accurate assessment of climate and land use impacts on regional hydrology.

2.8 <u>Summary</u>

The Community Climate Modeling System (CCSM3) sensitivity experiments are performed to investigate the impact of past wetland drainage on the hydro-climatology of Midwestern United States (MW USA). Coupled land surface and atmospheric components of CCSM3 are used at T85 (~140 km) horizontal grid mesh size to create four control model experiments. These include: (i) 355 ppm CO_2 for year 1990 excluding wetland (present condition), (ii) 355 ppm CO_2 for 1990 including wetland, (iii) 289 ppm CO_2 for year 1870 excluding wetland, and (iv) 289 ppm CO_2 for year 1870 including wetland. The CCSM3 control run for the present condition is validated with high resolution North American Regional Reanalysis (NARR) data for the Midwest region. Validation results show that CCSM3 is reasonable in simulating at-surface incoming solar radiation, net short wave radiation, and 2m air temperature. However, partitioning of net radiation into sensible and latent heat fluxes is imprecise, and summer precipitation is largely underestimated in CCSM3. To remove any biases in CCSM3 output, results from sensitivity experiments are analyzed in terms of difference in monthly climatology. Sensitivity experiment results show significant changes in summer sensible and latent heat fluxes due to wetland drainage. Near surface (2m) air temperature has significantly increased, and convective precipitation has decreased by a small amount (~50 mm) during the summer. Except for 2m air temperature which is affected by both green house gas emission based climate change and wetland drainage over the last century, all other climatic variables are primarily affected by wetland drainage in the region.

CHAPTER3. THE WATER AND ENERGY BUDGET STUDY

3.1 Introduction

Studies involving water and energy balances together provide two levels of constraints (closure of water and energy balance equations), and hence can lead to better conceptualization of the hydrologic system at basin scales. As a part of the Global Water and Energy Cycle Experiment (GEWEX), many studies were conducted with the goal of '*closing*' the water and energy balances for continental scale basins (Roads et al., 2002). Roads et al. (2003), hereafter referred as WEBS (Water and Energy Budget Synthesis), have synthesized water and energy budgets for the Mississippi River Basin (MRB) from the *best available models and observations* for the period 1996 – 1999. The WEBS study found that while model outputs qualitatively correspond with the available observations, large quantitative uncertainty exists among different model outputs. The limited number of tower flux observations (total of two) was cited as the rarest observations in the WEBS study. Since 2000, major developments have occurred with respect to improvements in regional reanalysis data (e.g. NARR; Mesinger et al., 2006), land surface modeling (e.g. CLM 3.5; Oleson et al., 2008), greater availability of energy flux observations (e.g. AmeriFlux data; Law et al., 2009), and availability of new land cover change datasets (Fry et al., 2009). Hence, it is worthwhile to revisit the WEBS study or a portion of that, and provide updated information about water and energy budgets in the MRB. As a part of a broader objective of assessing the impacts of climate and land cover changes on the water availability, this study presents an assessment of the reanalysis data and climate model outputs for quantifying water and energy budgets in MRB.

Energy and water fluxes within a hydrologic boundary/basin are linked through evapotranspiration (ET), which is a major component of the hydrologic cycle (Postel et al., 1996). Changes in ET brought by major land cover change can significantly impact regional fresh water availability, as well as regional ecosystems (Gordon et al., 2003; Zhang and Schilling, 2006). Despite the importance of ET, relatively few reliable estimates of ET are available compared to

runoff. Limited availability of observed ET is a major constraint for studying ET variability, and using ET for model evaluation or validation purposes in hydroclimatic studies. In the last decade, coordinated efforts have been made to measure carbon and water fluxes to assess changes in the terrestrial ecosystem (FLUXNET; Running et al., 1999; Baldocchi et al., 2001). FLUXNET provides a global network of over 500 flux measurement sites spreading across diverse biomes and climatic regions (http://daac.ornl.gov/FLUXNET/). FLUXNET coordinates among regional networks to ensure consistency and inter-comparability of the flux measurements, provide infrastructure support for data archival and distribution, and support discussion and synthesis of scientific results, with the overarching goal of providing validation datasets for net primary productivity, evaporation, and energy absorption at global scale. Hence, FLUXNET data provides an opportunity to improve our understanding of land surface and atmospheric interaction. AmeriFlux is the regional network of FLUXNET sites in America, and it provides a relatively denser network of observation sites in USA (Fig. 3.1). Thus, AmeriFlux data can be used to study the spatial and temporal pattern of ET, and for evaluating the performance of land surface hydrology models in MRB. The AmeriFlux data is available for a relatively short period (average data length: 6 years in this study) and only a few randomly distributed stations are available (total 16 in MRB), hence alternative sources of information need to be explored for large scale hydro-climatic studies.

Reanalysis data provide spatially and temporally continuous outputs for different surface and atmospheric variables by assimilating available observations from various sources (e.g. satellite data, meteorological observations from surface stations, and data from rawinsondes and dropsondes) with the help of atmospheric and land surface models. The North American Regional Reanalysis (NARR) data is a much improved version of reanalysis outputs compared to Global Reanalysis 1 and 2 (REAN1 – Kalnay et al., 1996; REAN2 – Kanamitsu et al., 2002) for hydroclimatic studies in the region (Mesinger et al., 2006). NARR outputs have been used to: (i) evaluate the performance of global and regional climate model (Kumar et al., 2010; Diffenbaugh, 2009), (ii) study the pattern of major hydroclimatic variability (e.g. precipitation recycling; Dominguez and Kumar, 2008; Dominguez et al., 2008), and (iii) assess the impacts of land use land cover change (Fall et al., 2009). However, reanalysis outputs can have biases and uncertainties, and the quality of outputs can vary depending upon the variable of interest (Maurer et al., 2001). This study evaluates the performance of NARR outputs for surface water and energy fluxes in MRB using independent observations and/or other model outputs.
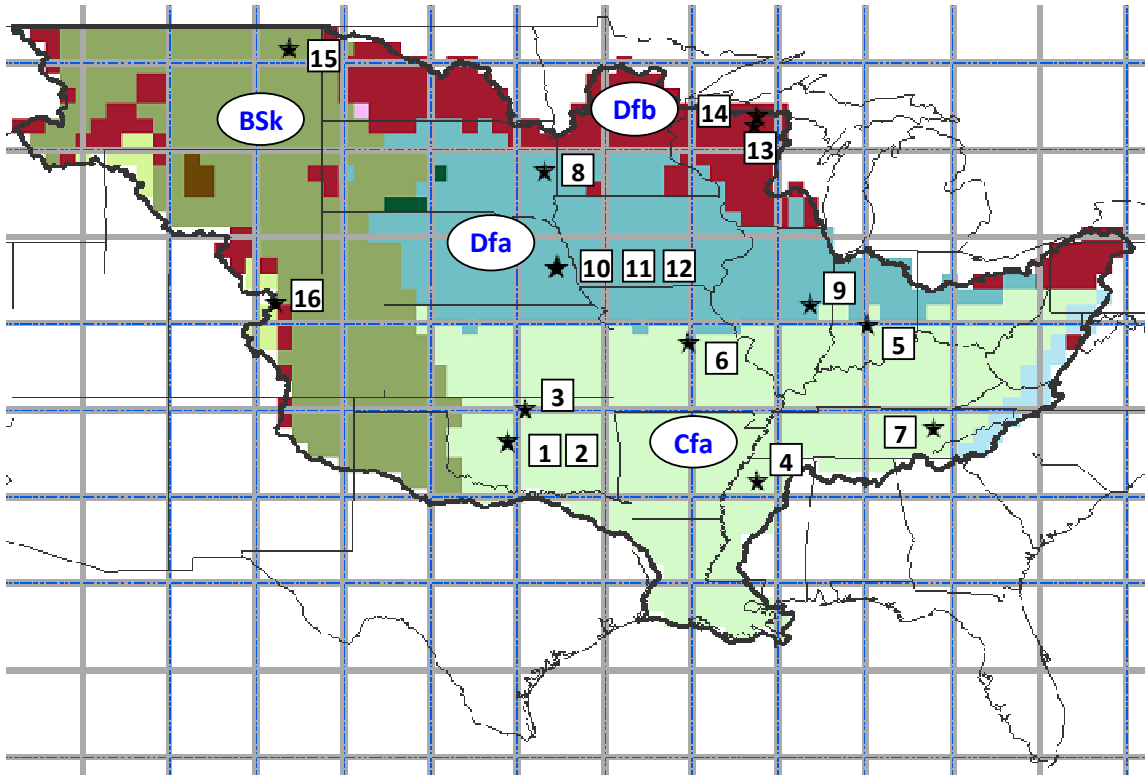
Fig. 3.1: The major climatic regions (Cfa, Dfa, Dfb, and BSk) and AmeriFlux stations in the Mississippi River Basin with CLM grid (T42 resolution) in the background. Climate classification is based on Köppen-Geiger climate classification map as described in Table (2). Each asterisk (★) represents the location of an AmeriFlux station. Geographical details of AmeriFlux stations are given in Table 3.1 according to the numbers shown on this map.

Understanding the evapotranspiration and precipitation feedback mechanism between climate change and land use land cover change is a critical component for the assessment of present and future water availability. The currently available coupled land surface and atmospheric modeling system (e.g. Community Climate System Model, Collins et al., 2006) provides an important tool to incorporate the feedback between land cover and climate. The coupled model outputs for surface water and energy fluxes can have biases, a portion of which can be attributed to biases in the atmospheric forcing (Lawrence et al., 2007). Hence, performance of coupled modeling system for surface water and energy fluxes should be evaluated using offline simulation of the land surface component model. Recently, model parameterization of the land component of the Community Climate System Model (CCSM) has been significantly improved with respect to ET partitioning, runoff scheme, ground water model, and frozen soil scheme (CLM3.5, Oleson et al,

2008; Stockli et al, 2008). This study evaluates the performance of CLM3.5 (hereafter referred as CLM) for surface water and energy fluxes in MRB.

Updated sources of water and energy flux observations, as well as reanalysis and climate model outputs are presented in the above discussion. These sources also have limitations such as: (i) point scale measurements from AmeriFlux sites; (ii) limited assimilation of surface observations in NARR (e.g., precipitation is assimilated in NARR, but ET and runoff are not); (iii) coarse resolution of CLM; and (iv) surface energy flux formulation and parameterization difference between NARR and CLM. This study lays out the foundation for accomplishing the broader objective of assessing the impacts of climate and land cover changes on the water availability in MRB by identifying different sources of uncertainties in the reanalysis and climate model outputs for water resources assessment. This study also provides an assessment of our progress in closing the water and energy balance in MRB in the last 10 years since the WEBS study.

3.2 Study Area, Data and Model Outputs

The Mississippi River Basin (MRB) is the largest river basin in North America, with more available observed data than any other major basin in the world (Roads et al., 2003). The MRB covers 41% of the conterminous USA, and has a total basin area of 3.2 million km^2. Major climatic gradients (temperature and precipitation) based on PRISM (Parameter-elevation Regressions on Independent Slopes Model, Daly et al., 1997, 1998) present day climate normal (1971-2000) are shown in Fig.3.2 (a) and (b). Basin average annual temperature and precipitation are 10.4 ^{O}C (range: -6.2 ^{O}C to 22 ^{O}C) and 810 mm (range: 144 mm/year to 2901mm/year), respectively. Annual average temperature shows a north to south gradient, and annual precipitation shows a north-west to south-east gradient (Fig. 3.2 (a) and (b)).

Major land cover classes in MRB based on the National Land Cover Data 2001 (NLCD 2001, http://www.epa.gov/mrlc/nlcd-2001.html) are shown in Fig. 3.2(c). NLCD 2001 provides high resolution (30 m) land cover data for the United States using satellite imagery (Landsat 5 and 7), and ancillary data (e.g. DEM, slope, aspect, population density) based on a decision tree classification algorithm, a supervised classification method (Homer et al, 2004; Breiman et al., 1984). Agriculture is the dominant land cover in MRB (39.2%), followed by grassland and shrubs (31.0%), forest (20.7%), wetland (3.6%), urban land cover (3.2%), open water (1.9%) and barren

(0.4%). Land cover change aggregated over local watershed scale (8-digit Hydrologic Unit Code (HUC8); http://nwis.waterdata.usgs.gov/tutorial/huc_def.html) based on NLCD 1992-2001 Land Cover Change Retrofit Product (Fry et al, 2009) is shown in Fig. 3.2(d). There are 851 HUC8 units in MRB, ranging in drainage area from 31 km^2 to 17287 km^2 (average area = 3826 km^2). Eighty HUC8 (9.4%) have experienced greater than or equal to 5% land cover change, and many of those HUC8s are located in the southern part of the basin (Arkansas-White-Red, and Lower Mississippi Basin; Figs. 3.2(c) and (d)). Overall, 2.5% of MRB has undergone land cover change between 1992 and 2001. The most dominant land cover transition includes a decrease in the forest area, and an increase in grassland / shrubs and urban area (Fig. 3.3).
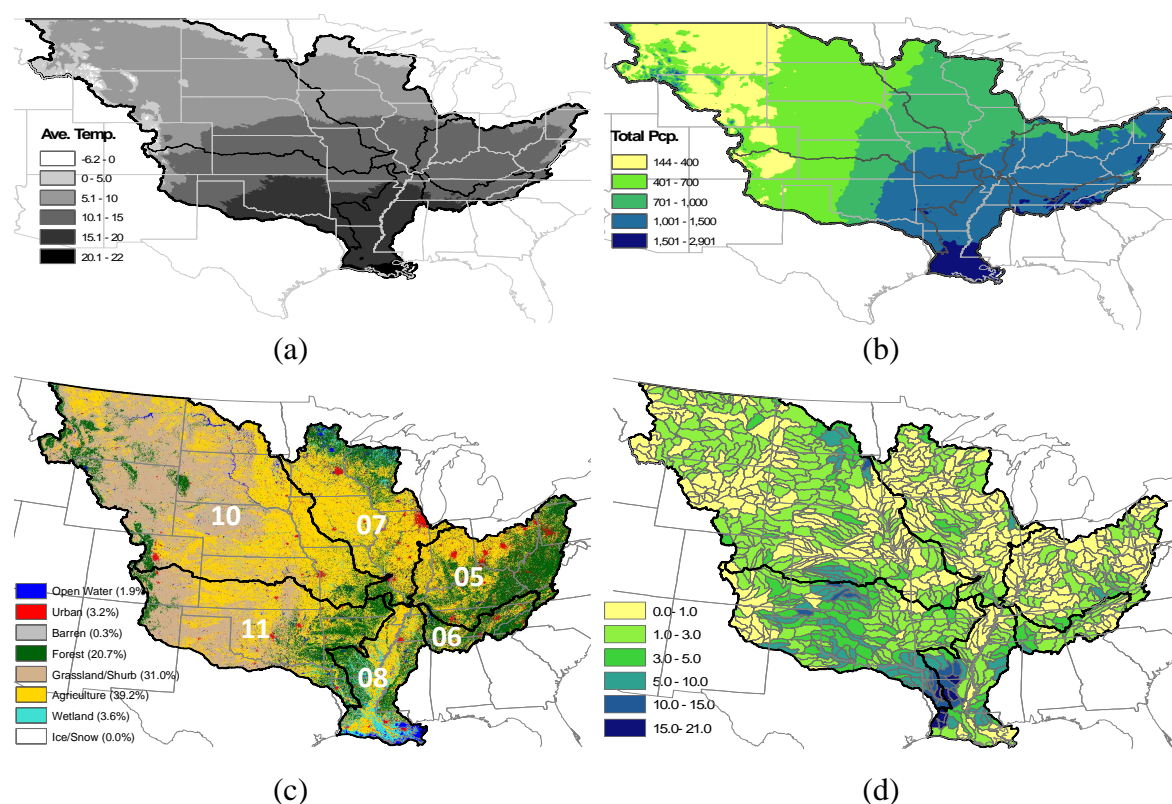
Fig. 3.2: Map of Mississippi River Basin showing: (a) annual average temperature (^{O}C); (b) annual total precipitation (mm/year) [data source: PRISM climate -normal 1971-2000]; (c) major land cover types (NLCD 2001); and (d) land cover change % (eight digit HUC watershed average, 1992 to 2001). Numbers in (c) correspond to major river basins/ water resource region in the MRB; 05: Ohio, 06: Tennessee, 07: Upper Mississippi, 08: Lower Mississippi, 10: Missouri, 11: Arkansas-White-Red.

Elevation in MRB ranges from sea level (0 m) at the mouth of the Mississippi river to 4282 m in the Rocky Mountains along the western boundary of the basin. Available sources of soil

characteristics (physical and hydraulic properties) information include SSURGO (The Soil Survey and Geographic) and STATSGO (The State Soil and Geographic) soil maps from the United States Department of Agriculture (USDA; http://soildatamart.nrcs.usda.gov/), and soil map provided by Miller and White (1998). More details on topography and soil characteristics can be found in the WEBS study.
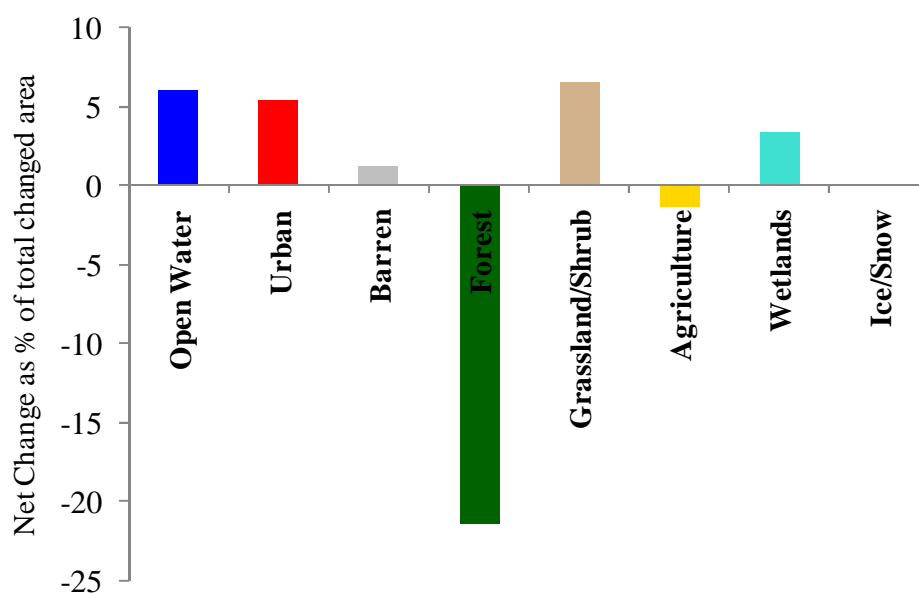
Fig. 3.3: Net land cover change between 1992 and 2001 in MRB

3.2.1   AmeriFlux Observations

The AmeriFlux Network was established in 1996, and it provides measurements of carbon, water, and energy fluxes in major vegetation types across different ecologic and climatic conditions in the Americas. Each flux tower represents an average footprint of one km radius at respective tower site (Baldocchi et al., 2001; Running et al., 1999). A total of 16 AmeriFlux sites are available in MRB (Fig. 3.1 and Table 3.1) with an average record length of 5.8 years (range: 2 – 12 years) between 1995 and 2007 (Law et al., 2009). The number of available flux tower observations has increased in the recent years, with most observations (86%) available since 1999. Monthly average Level 4 datasets of energy flux observations are included in this study (http://public.ornl.gov/ameriflux/level4data.html). In the Level 4 dataset, missing values for half hourly flux observations are filled with observations under similar meteorological conditions, and

night time fluxes are corrected for violation of eddy covariance method assumptions using u*
filtering (http://www.bgc-jena.mpg.de/bgc-mdi/html/eddyproc/index.html). All AmeriFlux sites
used in this study are referred using the following abbreviation: XX_YYY_ZZZ, where XX is the
country name (US for USA), YYY is the abbreviation of the site name, and ZZZ is the land cover
type. For example US_MMS_DBF refers to a "US" site named Morgan Monroe State (MMS) site
with a deciduous broad leaf forest (DBF) land cover type. A list of AmeriFlux sites included in
the analysis is provided in Table 3.1.

3.2.2   NARR

NARR provides a spatially continuous, high resolution (3-hourly temporal resolution, and 32-km
spatial resolution) regional reanalysis dataset for the North American domain since 1979. The
NARR dataset is developed as a major improvement upon the earlier National Center for
Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) Global
Reanalysis (REAN1) dataset in terms of resolution and accuracy (Mesinger et al. 2006). The
atmospheric component of NARR uses NCEP regional Eta model (Berbery et al, 2003; Mesinger
2000) with lateral boundary conditions from Global Reanalysis-2 (REAN2; Kanamitsu et al.,
2002) and the Eta data assimilation system (Rogers et al., 1996). The land component of NARR
uses the Noah land surface model (Ek et al., 2003; Chen and Dudhia, 2001). Major observations
assimilated in NARR include: (i) precipitation data from rain gauging stations; (ii) radiance data
from satellite observations; (iii) near surface wind (10m) and moisture (2m) data from Global
Reanalysis-2 outputs; (iv) sea and lake surface temperature; and (v) sea and lake ice cover data.
Successful assimilation of high quality detailed precipitation observations in NARR provides an
improved dataset for studying land surface hydrology (e.g. soil moisture), and land atmospheric
interactions (Mesinger et al., 2006).

Table 3.1: List of AmeriFlux Sites.

Sl. No	AmeriFluxSite	State	Site Name	Lat. (degree North)	Long. (degree East)	Site Elevation (m)	NARR Elevation (m)	CLM Elevation (m)	NARR Land Cover	CLM Land Cover
										FR/ShGr/CRO
1	US_ARb_GRA	OK	ARM SGP burn site-Lamont	35.55	-98.04	424	487	400	CRO	19/57/1
2	US_ARc_GRA	OK	ARM SGP control site- Lamont	35.54	-98.04	424	487	400	CRO	19/57/1
3	US_ARM_CRO	OK	ARM SGP site-Lamont	36.61	-97.49	414	379	503	CRO	4/30/52
4	US_Goo_GRA	MS	Goodwin Creek	34.25	-89.97	87	63	87	DBF	24/20/56
5	US_MMS_DBF	IN	Morgan Monroe State Forest	39.32	-86.41	275	188	214	CRO	10/10/79
6	US_Moz_DBF	MO	Missouri Ozark Site	38.74	-92.20	219	186	299	DBF	29/40/31
7	US_WBW_DBF	TN	Walker Branch Watershed	35.96	-84.29	343	379	412	DBF	52/33/14
8	US_Bkg_GRA	SD	Brookings	44.35	-96.84	510	541	409	CRO	3/2/95
9	US_Bo1_CRO	IL	Bondville	40.01	-88.29	219	186	214	CRO	10/10/79
10	US_Ne1_CRO	NE	Mead - irrigated maize site	41.17	-96.48	361	379	351	CRO	5/15/78

Table 3.1: List of AmeriFlux Sites (continued)

Sl. N o	AmeriFluxSit e	State	Site Name	Lat. (degree North)	Long. (degree East)	Site Elevatio n (m)	NARR Elevatio n (m)	CLM Elevati on (m)	NARR Land Cover	CLM Land Cover FR/ShGr/ CRO
12	US_Ne3_CRO	NE	Mead - rainfed maize-soybean	41.18	-96.44	363	379	351	CRO	5/15/78
13	US_WCr_DBF	WI	Willow Creek	45.81	-90.08	515	487	363	MF	41/16/20
14	US_Los_CSH	WI	Lost Creek	46.08	-89.98	480	487	363	MF	41/16/20
15	US_FPe_GRA	MT	Fort Peck	48.31	-105.10	634	708	670	BSH	2/22/56
16	US_NR1_ENF	CO	Niwot Ridge Forest (LTER NWT1)	40.03	-105.55	3050	2976	2531	ENF	19/57/1

NARR and CLM Elevation and Land Cover represent nearest grid point elevation and land cover in the respective dataset. CLM land cover is grouped under three categories % Forest (FR), %Shrub and Grass (ShGr) and % Crop (CRO). AmeriFlux sites are referred as XX_YYY_ZZZ, where XX is the country name (US for USA), YYY is the abbreviation of the site name, and ZZZ is the land cover type (CRO- Crop, CSH – Closed Shrubland, DBF – Deciduous Broadleaf Forest, ENF- Evergreen Needleleaf Forest, GRA – Grassland). ARM – Atmospheric Radiation Measurement, SGP – Southern Great Plains, MF – Mixed Forest, BSH – Broadleaf Shrub with ground cover. The state abbreviations are CO: Colorado, IL: Illinois, IN: Indiana, MO: Missouri, MS: Mississippi, MT: Montana, NE: Nebraska, OK: Oklahoma, SD: South Dakota, TN: Tennessee, WI: Wisconsin

3.2.3   CLM offline simulation

CLM (version 3.5) is the recently released land component of the Community Climate System Model (Collins et al., 2006; Oleson et al., 2008). Major hydrological processes in CLM include canopy interception, transpiration, through fall, evaporation, infiltration, surface and subsurface runoff, and water table dynamics. A grid cell is first divided into four major land units (vegetative cover, lake, wetlands and glacier), and the vegetative fraction of the grid cell can have a maximum of four Plant Functional Types (PFT) out of a total 16 PFTs (Oleson et al., 2004). Surface data into CLM (e.g. PFTs, leaf and stem area) is based on multiyear MODIS land surface data at 0.5^0 resolution (Lawrence and Chase, 2007). Several improvements in land surface parameterization have been incorporated in CLM to alleviate water and energy biases observed in its predecessor CLM3 (Oleson et al., 2008; Stockli et al, 2008). Improvements include an improved canopy evaporation scheme, simple based surface and subsurface runoff scheme based on the distributed hydrologic model called TOPMODEL (Niu et al., 2005; Niu and Yang, 2006), simple ground water model (Niu et al., 2007), and a new frozen soil scheme (Oleson et al., 2008). Offline results of CLM provided by Oleson et al. (2008) are used in this study. The offline simulation of CLM uses atmospheric forcing data from Qian et al. (2006) for 1948 – 2004, and has a long spin up period (624 years), by cycling the same atmospheric forcing (1948-2004) 12 times, to stabilize the deep soil water in the model. Detailed description of the CLM model, simulation, and results are provided in Oleson et al. (2008). The atmospheric forcing was constructed by adjusting REAN1 outputs using observed monthly precipitation and temperature, satellite radiation data, and cloud cover data (Qian et al., 2006). Because of model design (water and energy balance closure for each grid cell and each time step), and gridded observational input data set, CLM is expected to show better results for surface water and energy fluxes in MRB.

3.2.4   Other Dataset and Models

The United States Historical Climatology Network (USHCN) data from 71 stations in Indiana and Illinois are used to study the characteristics of spatial scale of climate forcing in the region. Monthly time series of precipitation and temperature for 113 years (1896 to 2008) are included in this study. USHCN data incorporates adjustment for observation biases (Karl et al. 1986; Vose et al. 2003), and artificial changes in the time series arising due to station relocation and equipment

change (Menne et al. 2009). Because gridded or denser network of tower flux observations are not available, long term observed records of precipitation and temperature are used to supplement the analysis of point scale measurements versus climate model grid cells.

Three runoff datasets are also used in this study including: (i) naturalized runoff estimates for MRB from Maurer and Lettenmaier (2001), (ii) University of New Hampshire / Global Runoff Data Centre (UNH- GRDC runoff data; Fekete et al., 2002), and (iii) runoff data from the Variable Infiltration Capacity (VIC) model (Maurer et al., 2002). The naturalized runoff data in Maurer and Lettenmaier (2001) were created by adding consumptive water use to the observed runoff with the help of observed consumptive water use statistics and VIC model output. Consumptive water use accounts for 6% of the naturalized runoff during the 1988 to 1999 period (range: 4 to 7%). UNH - GRDC runoff were created by combining the Water Balance Model outputs with the observed mean annual runoff data (Fekete et al, 2002).

PRISM monthly precipitation and temperature data (1980 to 2004) are also used in this study as climate observations. PRISM is a high quality spatial dataset at 4 km resolution created by using point observations of precipitation and temperature with Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 1997; 1998).

3.3 Methodology

The methodology involves: (i) regional classification of the study area (MRB) based on the climatic condition; (ii) re-gridding of NARR (NARR_Regrid) to the climate model grid size; (iii) long term (25 years) climatology comparison between NARR, NARR_Regrid, and CLM; (iv) evaluation of NARR, NARR_Regrid, and CLM with respect to AmeriFlux observations using equal sample size principle; (v) evaluation of the spatial and temporal variability in total runoff in MRB; and (vi) analysis of water and energy balance closure in MRB. Each step in the methodology is briefly described below.

3.3.1   Regional Classification of MRB

The four major climatic regions in MRB (Fig. 1) are: (i) Cfa –Warm temperate climate, fully humid, and hot summer; (ii) Dfa – Snow Climate, fully humid, and hot summer; (iii) Dfb – Snow Climate, fully humid, and warm summer; and (iv) BSk – Arid, cold steppe climate. The regional classification shown in Fig. 1 is based on the digital Köppen-Geiger climate classification map provided by Kottek et al. (2006). For climate classification, Kottek et al. (2006) have used 0.5^O resolution monthly temperature and precipitation data from Climate Research Unit (www.cru.uea.ac.uk), and Global Precipitation Climatology Center (http://gpcc.dwd.de), respectively for 1951 to 2000. Major climatic characteristics of the four regions are listed in Table 3.2. South-East MRB has temperate climate, North-East MRB has snow climate, and Western MRB has arid climate.

Table 3.2: Major climatic regions in MRB (Kottek et al., 2006).

Sl. No.	Region	T_{min}	T_{max}	Precipitation
1	Cfa	$-3^O C < T_{min} < 18\ ^O C$	$T_{max} > 22\ ^O C$	neither dry summer nor dry winter
2	Dfa	$T_{min} < -3\ ^O C$	$T_{max} > 22\ ^O C$	neither dry summer nor dry winter
3	Dfb	$T_{min} < -3\ ^O C$	$T_{max} < 22\ ^O C$ and $4\ T_{mon} \geq 10\ ^O C$	neither dry summer nor dry winter
4	BSk	$T_{ann} < 18\ ^O C$		$5\ P^{th} < P_{ann} < 10\ P^{th}$

T_{min}: Minimum monthly mean temperature; T_{max}: Maximum monthly mean temperature; T_{ann}: Annual mean temperature; T_{mon}: monthly mean temperature. Dry summer and dry winter are defined as a function of monthly total precipitation during the summer (May-Oct.) and winter (Nov. – April) months. For dry summer: $P_{smin} < P_{wmin}$, $P_{wmax} > 3\ P_{smin}$ and $P_{smin} < 40$ mm; and for dry winter $P_{wmin} < P_{smin}$ and $P_{smax} > 10\ P_{wmin}$. Where, P_{smin}, P_{smax}, P_{wmin} and P_{wmax} are minimum and maximum monthly total precipitation values during the summer and winter months, respectively. P^{th} (mm) is the dryness threshold for the arid region which is a function of annual average temperature. $P^{th} = (2* T_{ann})$ if 2/3 of annual precipitation occurs in winter; $(2* T_{ann} + 28)$ if 2/3 of annual precipitation occurs in summer; $(2* T_{ann} + 14)$ otherwise. P_{ann} is annual total precipitation (mm/year)

56

3.3.2   Regridding of NARR

One of the objectives of this study is to evaluate the prediction of ET by the land component (CLM) of a global climate model at regional scale. Because NARR's spatial resolution is notably higher (32 km) compared to CLM resolution (T42, ~ 280 km), NARR outputs are re-gridded to CLM resolution in a two step process using the National Center for Atmospheric Research's (NCAR) Command Language (NCL; http://www.ncl.ucar.edu/). In the first step, NARR outputs are regridded to 0.5 degree (~50 km) resolution using inverse distance weighting, and in the second step, 0.5 degree output is then re-gridded to T42 resolution using area average method. This two step procedure is followed because NARR has Lambert conformal conic native grid projection. NARR outputs regridded to T42 resolution are referred as NARR_Regrid in this study.

3.3.3   Monthly Climatology Comparison

Monthly climatology (long term monthly mean and inter-annual variability) of near surface hydroclimatic variables is prepared from NARR, NARR_Regrid, and CLM monthly outputs from 1980 to 2004 (25 years). The monthly climatology is presented with 95% confidence interval uncertainty range calculated from standard deviations of 25 years monthly outputs in each case. The uncertainty range represents the inter-annual variability during the analysis period. Because AmeriFlux data are not available for the 25 year period, these data are not included in the 25 years monthly climatology comparison.

3.3.4   CLM, NARR, and AmeriFlux Comparison

Monthly averages of CLM, NARR, and AmeriFlux observations are compared at 16 AmeriFlux site locations in MRB. Because the spatial and temporal coverage of AmeriFlux observations is not consistent with that of CLM and NARR, the equal sample size principle (Robock et al., 2003) is used for making this comparison. In the equal sample size principle, point observations at a site (AmeriFlux site in this study) are compared with the model outputs (CLM and NARR) from the nearest grid cell for the available time period of observation.

3.3.5   Closing Water and Energy Balance for MRB

Water and energy balance components are linked through ET as given in Eqs. 3.1-3.3 below.

$$\Delta s = P - ET - N \qquad\qquad (3.1)$$

$$R_n = Sht + Lht + Ght \qquad\qquad (3.2)$$

$$Lht = \lambda * ET \qquad\qquad (3.3)$$

Where Δs is the change in storage (in the active soil layers), P is precipitation, ET is evapotranspiration, N is total runoff, R_n is net radiation including short and long wave radiation, Sht is sensible heat flux, Lht is latent heat flux, Ght is ground heat flux, and λ is the latent heat of vaporization. Lht and ET are used interchangeably in this study. Summer is considered to be from May to October (6 months) and winter is considered to be from November to April (6 months), unless specified otherwise.

Any significant bias in ET will reflect bias in the estimation of runoff from the basin, because P is the constrained term in Eq. 3.1 (precipitation is observed forcing data in CLM, and precipitation observation is assimilated in NARR) and Δs can be taken as zero for long term annual water balance. Spatial distribution of runoff is validated with the UNH-GRDC runoff (Fekete et al., 2002) and VIC runoff (Maurer et al., 2002). To exclude the effect of water withdrawals for irrigation or water supply for cities, total runoff at the watershed outlet is compared with the naturalized runoff estimate from Maurer and Lettenmaier (2001).

3.3.6   Statistical Methods

Comparison between different datasets and models are done using the following statistical measures: mean, standard deviation (inter-annual variability), bias (model – observation), pearson product moment correlation coefficient (correlation coefficient), square of correlation coefficient (R^2), root mean square error (RMSE), semivariogram (plot of 0.5 * [square difference] against separation distance, [Kitanidis, 1997; page 32-40]), and statistical significance of difference in monthly mean values. For statistical significance T-test is used with 95% confidence

interval ($\alpha = 0.05$). Ninety five percent uncertainty range for long term monthly mean value (μ) is calculated by using Eq. 3.4 (Miller and Miller 2004, page 358).

$$\bar{x} - t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}, n-1} \cdot \frac{s}{\sqrt{n}} \qquad (3.4)$$

where $\bar{x}$ and s are mean and standard deviation of random sample of size n ($= 25$) from a normal population, α is the significance level ($= 0.05$), and values of $t_{\frac{\alpha}{2}, n-1}$ is taken from T-distribution table. The assumption of normality for all variables (basin average each month time series from 1980-2004) is checked using statistical test (Shapiro-Wilk test) and graphical method (Quantile-Quantile plot) in SAS and found valid.

3.4 <u>Results</u>

3.4.1   <u>Monthly Climatology Comparison</u>

Basin average climatological mean and variability (1980-2004) for water and energy balance components in MRB are presented in Table 3.3. Mean annual net radiation obtained from CLM (69.7 W/m^2) is lower (18%) compared to NARR (84.7 W/m^2). Sensible and latent heat flux are also lower by 19% and 31%, respectively in CLM compared to NARR. Latent heat flux represents 59% of the net radiation in CLM, and it represents 70% of the net radiation in NARR. Ground heat flux is a minor component of the energy balance equation (Eq.3.2). Higher inter-annual variability (2.2 W/m^2) in ground heat flux in comparison to the overall mean (0.5 W/m^2) is due to the opposing sign of ground heat flux during summer and winter months. CLM does not produce any closing error in the energy balance equation (*Rn – Lht – Sht - Ght = 0.0*), but NARR shows an error of -9.6 W/m^2 (11% of net radiation) in the energy balance equation. Regridding of NARR (NARR_Regrid) has resulted in less than 1% reduction in the basin average monthly mean of net radiation, sensible and latent heat fluxes, in comparison to original NARR outputs. However, inter-annual variability (standard deviation) is reduced by 9% for net short wave, 6% for net long wave, 21% for sensible heat flux, and 20% for latent heat flux as a result of regridding NARR outputs (Table 3.3).

Table 3.3: Climatological annual mean (1980-2004) for the MRB from CLM and NARR outputs.

Variable	Unit	PRISM	NARR	CLM	NARR_Regrid
Net short wave (+ down)	W/m^2	NA	159.2(11.1)	132.2(6.9)	158.9(10.1)
Net long wave (+ up)	W/m^2	NA	74.5(9.1)	62.5(4.1)	74.5(8.6)
Sensible heat flux	W/m^2	NA	34.1(11.1)	27.5(6.0)	33.9(8.8)
Latent heat flux	W/m^2	NA	59.7(11.2)	41.4(5.2)	59.1(8.9)
Ground heat flux	W/m^2	NA	0.5(2.2)	0.7(3.4)	0.5(2.1)
2 m air Temperature	^{O}C	10.5(2.0)	11.4(2.0)	10.7(1.7)	11.2(2.0)
Precipitation (P)	mm/year	806(446)	751(420)	792(328)	736(355)
Total Runoff (N)	mm/year	NA	91(107)	269(143)	91(71)
ET/P	NA	NA	1.00	0.66	1.01
N/P	NA	NA	0.12	0.34	0.12

Numbers in parenthesis represent average values of monthly standard deviation. For precipitation and runoff monthly standard deviation is multiplied by 12 to be consistent with annual total value. PRISM climatological mean is also presented for the reference purpose.

Twenty five years (1980-2004) monthly climatology of near surface air temperature (T_{air}) and precipitation (*P*) from PRISM, NARR, CLM, and NARR_Regrid are shown in Fig. 3.4. In comparison to PRISM data, basin average mean annual temperature is 0.9 ^{O}C higher in NARR, 0.2 ^{O}C higher in CLM and 0.7 ^{O}C higher in NARR_Regrid, and basin average annual total precipitation is 7% lower in NARR, 2% lower in CLM, and 9% lower in NARR_Regrid (Table 3.3). Regridding of NARR (NARR_Regrid) has resulted in 18% reduction in the inter-annual variability of precipitation compared to the original NARR outputs. There is no change in the inter-annual variability of temperature between NARR and NARR_Regrid. Inter-annual variability in PRISM precipitation data (446 mm / year) is closer to NARR (420 mm / year), and inter-annual variability in CLM precipitation data (328 mm/year) is closer to NARR_Regrid (355 mm / year). Basin average monthly precipitation from NARR, NARR_Regrid, and CLM are not statistically different (90% confidence interval) compared to PRISM precipitation data for all 12

months (not shown). Basin average monthly temperature from NARR and NARR_Regrid are statistically different (90% confidence interval) compared to PRISM temperature data during summer months (May to Oct), but they are not statistically different during winter months (not shown). For CLM, monthly temperature is statistically different compared to PRISM temperature data for three summer months (May to July), and they are not statistically different for the remaining nine months. Statistical difference in summer temperature can be due to higher magnitude of mean value and lower inter-annual variability. For example, summer temperature mean and standard deviations are 18.6 ^{O}C and 1.5 ^{O}C respectively; whereas, winter temperature mean and standard deviations are 2.4 ^{O}C and 2.5 ^{O}C respectively, for MRB in PRISM data.

Spatial variability of the absolute difference between CLM and NARR_Regrid annual average latent heat flux and sensible heat flux are presented in Fig. 3.5(a) and (b), respectively. For latent heat flux, the difference between CLM and NARR_Regrid shows an east-west divide. In the eastern part, CLM latent heat flux is lower compared to NARR_Regrid for all months (-38% annual average difference, Fig. 3.5 (c)). In the western part, opposite signs of difference in first (negative) and second half (positive) of the year cancel each other (Fig. 3.5 (d)), making the annual average difference smaller (-18%). In the eastern MRB, CLM sensible heat flux is higher in summer, and is lower in winter, compared to NARR_Regrid, making the annual average difference smaller (+16%; Fig. 3.5(e)). In the western MRB, CLM sensible heat flux is lower in summer compared to NARR_Regrid, and annual average difference is -38% (Fig. 3.5 (f)). As shown in Table 3.3, NARR_Regrid has resulted in less than 1% reduction in the mean annual values of sensible and latent heat fluxes. Therefore, the difference in NARR and CLM sensible and latent heat fluxes in the eastern and western parts of MRB should be similar to the difference between NARR_Regrid and CLM. Thus, NARR and CLM provide spatially (east versus west) and temporally (summer versus winter) different estimate of sensible and latent heat flux in MRB.

NARR has a 12% water balance closing error (*ET* + *N = 1.12 P*); whereas, CLM does not have water balance closing error (*ET* + *N = P)*. The water balance closing error in NARR is not affected by regridding of NARR outputs (Table 3.3). Comparison of NARR, NARR_Regrid, and CLM runoff outputs with observed data is presented in Section 3.4.4.
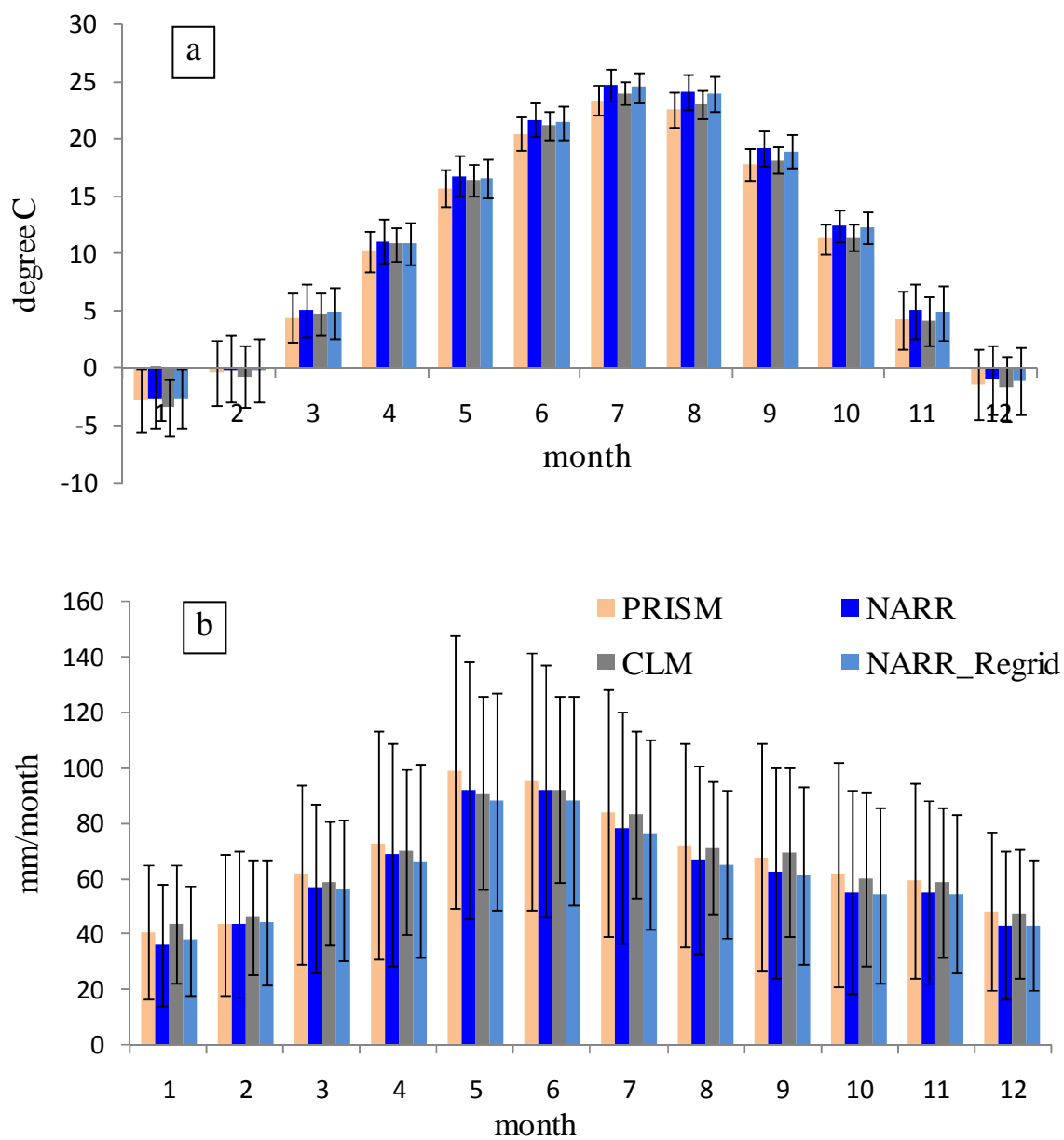
Fig. 3.4: Monthly climatology of basin average (a) 2m air temperature and (b) precipitation from PRISM, NARR, CLM, and NARR_Regrid in MRB. Monthly average values from 1980 to 2004 are shown in this figure, and error bars represent ±1 standard deviation of inter-annual variability from 1980 to 2004.
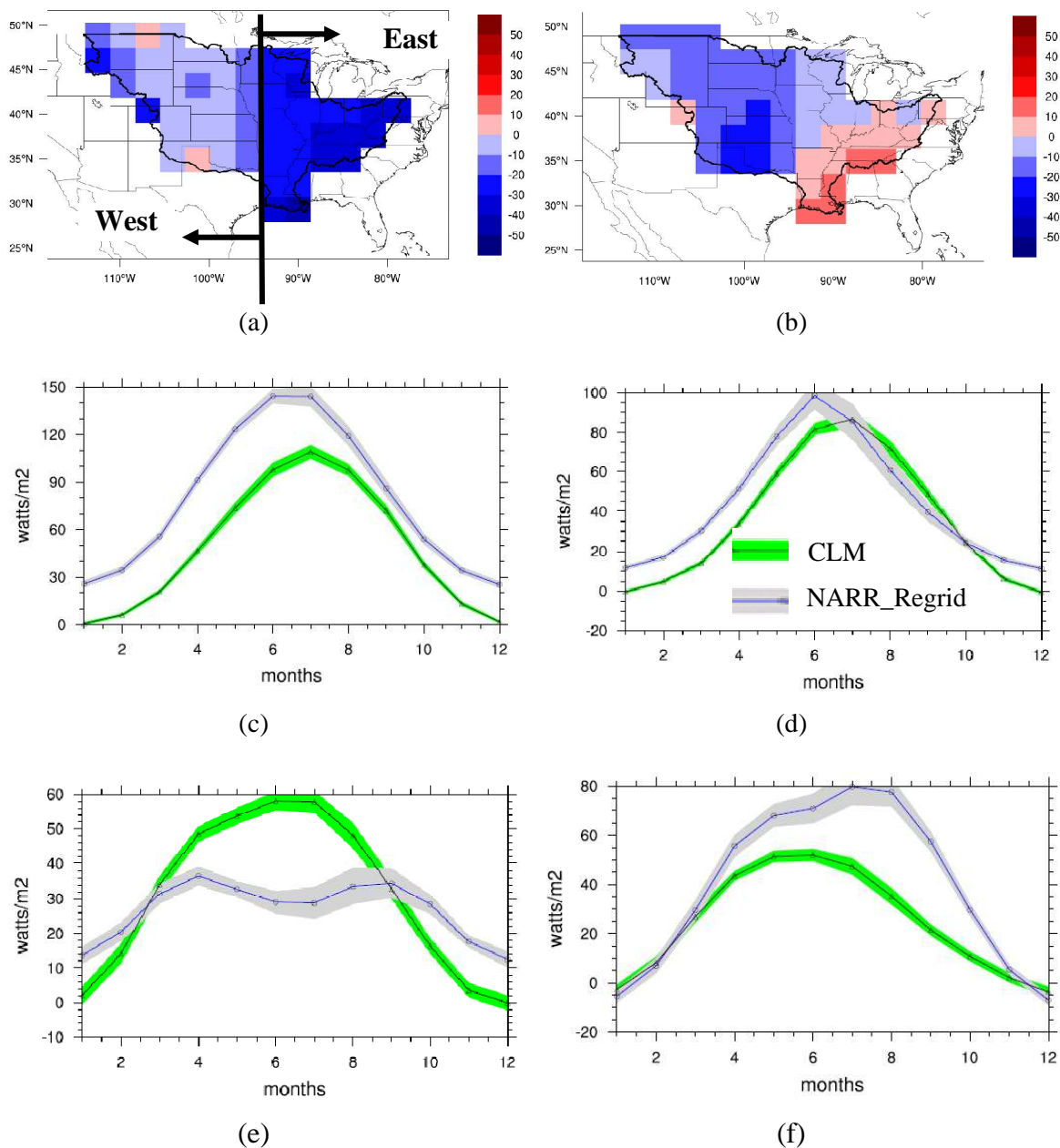
Fig. 3.5: Comparison of CLM and NARR_Regrid surface energy fluxes in MRB. (a) Absolute difference (CLM – NARR_Regrid) in W/m^2: (a) mean annual latent heat flux in W/m^2, and (b) mean annual sensible heat flux. Monthly climatology of latent heat flux in: (c) east, and (d) west part of MRB. Monthly climatology of sensible heat flux in: (e) east, and (f) west part of the MRB. Shaded region represent 95% uncertainty range. East and west portions of MRB are marked in Fig. (a)

3.4.2   AmeriFlux, NARR, and CLM Comparison

Point scale observations at 16 AmeriFlux sites in MRB are compared with the nearest grid cell in NARR and CLM using the equal sample size principle (Section 3.3.4). Monthly mean values of observed hydroclimatic variables at AmeriFlux sites are given in Table 3.4. For 12 AmeriFlux sites, results are presented until 2004 because CLM outputs are available from 1948 – 2004. For sites that have less than two years of data prior to 2004 (a total of four sites), comparison extends beyond 2004, and is made only with NARR outputs. Elevations and land cover types at AmeriFlux sites and the corresponding nearest grid cells in NARR and CLM data are given in Table 3.1.

The number of AmeriFlux sites present in Cfa, Dfa, Dfb, and BSk climate regions are 7, 5, 2 and 2, respectively. In Cfa region, three grassland sites have higher average latent heat flux to incoming solar radiation ratio (average = 0.30, range: 0.26 to 0.33), compared to three deciduous broadleaf forest sites (average: 0.27, range: 0.25 to 0.30), and one crop land site (0.23). In Dfa region, latent heat flux to incoming solar radiation ratio is higher at one grassland site (0.39) compared to four crop sites (average: 0.27, range: 0.26 to 0.28). In Dfb region, one deciduous broadleaf forest site and one closed shrub land site has the same latent heat flux to incoming solar radiation ratio (0.19). In Bsk region, one evergreen needleleaf forest site has higher latent heat flux to incoming solar radiation ratio (0.25) compared to one grassland site (0.14).

NARR and CLM monthly outputs are compared with AmeriFlux observations in terms of R^2, Bias (model – observation), and RMSE (Root Mean Square Error). Both Bias and RMSE are expressed as percentage of mean observations at respective sites (Table 3.5). AmeriFlux site US_NR1_ENF, located in Rocky mountain range (elevation – 3050m; Monson et al., 2005), shows large difference (2.6 $^\circ$C higher) in mean temperature compared to the nearest CLM grid cell. Hence, US_NR1_ENF results are not included in 11 sites average results presented in the next paragraph. Out of remaining 15 sites, only 11 sites are included in comparison with NARR because CLM outputs are not available at four AmeriFlux sites. These 11 sites are indicted in bold letters in Table 3.5.

Table 3.4: Statistical summary of the AmeriFlux monthly observations

Reg	Sl. No.	AmeriFluxSite	Time Period	Observation				
				Mean value for given Time Period				
				R_g	P	T_{air}	Lht	Sht
Cfa	1	US_Arb_GRA	2005-2006	192.7	60.9	16.8	50.9	47.2
	2	US_Arc_GRA	2005-2006	178.3	64.7	16.4	55.8	45.9
	3	US_ARM_CRO	2003-2004	180.1	64.3	14.7	40.9	29.5
	4	US_Goo_GRA	2002 - 2004	181.3		17.4	60.1	24.2
	5	US_MMS_DBF	1999 - 2004	167.0	87.4	12.1	41.5	22.2
	6	US_Moz_DBF	2004-2007	185.6	68.4	14.4	56.0	37.7
	7	US_WBW_DBF	1995-1999	172.2		14.8	45.1	30.4
Dfa	8	US_Bkg_GRA	2004-2006	172.0	56.1	7.9	66.6	16.7
	9	US_Bo1_CRO	1997-2004*	167.5	66.6	11.4	46.9	25.9
	10	US_Ne1_CRO	2002-2004	177.3	71.0	10.6	48.9	20.6
	11	US_Ne2_CRO	2002-2004	180.9	68.8	10.4	49.4	20.7
	12	US_Ne3_CRO	2002-2004	176.8	50.8	10.2	45.3	26.4
Dfb	13	US_WCr_DBF	1999 - 2004	145.9	63.2	5.3	27.3	24.4
	14	US_Los_CSH	2001-2004	159.3	57.8	5.4	29.6	23.0
BSk	15	US_Fpe_GRA	2000-2004	159.0	32.8	5.4	21.7	25.2
	16	US_NR1_ENF	1999-2004	182.7	54.4	2.2	46.0	42.4

[R_g – incoming solar radiation (W/m^2), P – Precipitation (mm/month), T_{air} - air Temperature (^{O}C), Lht – Latent Heat Flux (W/m^2), Sht – Sensible Heat Flux (W/m^2)]. Precipitation for US_Goo_GRA and US_WBW_DBF are not shown because of missing observations

Table 3.5: Model performance evaluation, Bias and RMSE are expressed as the % of observed mean values given in Table 3.4.

Sl. No.	NARR					CLM				
	R2/Bias/RMSE					R2/Bias/RMSE				
	R_g	P	T_{air}	Lht	Sht	R_g	P	T_{air}	Lht	Sht
1	0.80/29/36	0.82/3/31	0.96/2/14	0.64/9/47	0.01/11/91					
2	0.78/40/47	0.79/-3/33	0.96/5/15	0.63/-1/45	0.00/14/89					
3	**0.95/32/35**	**0.85/7/28**	**1.00/6/8**	**0.80/53/72**	**0.84/50/72**	**0.91/3/13**	**0.50/-9/50**	**1.00/-4/6**	**0.68/23/60**	**0.64/-22/62**
4	**0.83/20/26**		**0.87/1/17**	**0.74/56/70**	**0.08/2/49**	**0.86/4/15**		**0.87/0/17**	**0.67/-2/40**	**0.25/58/100**
5	**0.97/24/25**	**0.71/10/33**	**0.97/1/14**	**0.82/95/117**	**0.31/-4/64**	**0.94/-1/11**	**0.58/-7/36**	**0.96/-6/18**	**0.93/6/28**	**0.16/7/93**
6	0.97/20/22	0.65/13/44	0.95/4/15	0.78/41/60	0.47/-17/41					
7	**0.97/22/24**		**0.99/-3/6**	**0.84/89/104**	**0.36/-12/37**	**0.94/5/11**		**0.99/1/6**	**0.91/11/26**	**0.05/42/99**
8	0.97/23/26	0.90/-7/32	0.91/15/46	0.89/-25/34	0.40/106/191					
9	**0.95/24/26**	**0.61/23/51**	**1.00/5/7**	**0.89/55/80**	**0.64/-5/37**	**0.94/-3/11**	**0.41/21/55**	**0.99/1/8**	**0.85/-9/37**	**0.34/-10/67**
10	0.99/23/25	0.47/-25/70	1.00/3/11	0.42/8/78	0.02/86/191	0.96/0/8	0.45/-9/68	1.00/6/8	0.85/4/37	0.31/4/94
11	0.99/21/23	0.43/-22/76	0.98/6/16	0.37/7/79	0.02/85/197	0.96/-2/8	0.36/-7/77	0.99/9/14	0.79/3/43	0.33/4/100
12	0.99/24/26	0.88/5/29	0.99/7/15	0.51/17/71	0.22/45/117	0.96/1/8	0.72/27/52	0.99/11/14	0.84/12/43	0.46/-19/66
13	0.97/31/35	0.79/5/33	0.99/4/23	0.89/136/163	0.55/-14/52	0.96/4/12	0.58/-7/42	0.98/-11/32	0.91/32/58	0.56/-34/77
14	0.93/19/26	0.32/16/65	0.93/3/54	0.92/109/132	0.68/7/38	0.92/-5/14	0.29/-2/62	0.92/-13/59	0.91/18/47	0.61/-25/77
15	**0.98/26/29**	**0.27/-26/111**	**0.98/32/44**	**0.57/24/84**	**0.89/51/92**	**0.96/-1/11**	**0.23/1/109**	**0.99/-34/44**	**0.51/23/104**	**0.67/-29/63**
16	0.81/34/41	0.82/-10/38	0.97/-28/77	0.84/37/55	0.40/-81/102	0.81/3/18	0.45/-47/76	0.96/121/159	0.81/-42/48	0.76/-20/35

Comparison between NARR and CLM outputs show that incoming solar radiation (R_g) and temperature (T_{air}) are the two most correlated variables in these datasets. Average R^2 for R_g / T_{air} at 11 AmeriFlux sites is 0.96 / 0.97 in NARR, and 0.94 / 0.97 in CLM. Compared to AmeriFlux data, incoming solar radiation is 24% higher in NARR (range: 19% to 32%), and 0.5% higher in CLM (range: -5% to + 5%). Average near surface air temperature is 6% higher in NARR (range: -3% to 32%), and 3% lower in CLM (range: -34% to 11%). NARR outputs show higher correlation (average R^2: 0.59, range: 0.27 to 0.88) with observed precipitation compared to CLM (average R^2 :0.46, range:0.23 to 0.72). Average monthly precipitation is 1% lower in NARR (range: -26% to 23%), and 1% higher in CLM (range: -9% to 27%) [Precipitation results include comparison at 9 sites only, because at remaining 2 sites precipitation observation was not satisfactory because of missing values]. CLM outputs for latent heat flux show higher correlation (average R^2: 0.80, range: 0.51 to 0.93) with observation compared to NARR outputs (average R^2: 0.71, range: 0.37 to 0.92). Average monthly latent heat flux is 11% higher in CLM (range: -9% to 32%), and 59% higher in NARR (range: 7% to 136%). NARR outputs for sensible heat flux also show slightly higher correlation with observations (average R^2: 0.42, range: 0.02 to 0.89) compared to CLM outputs (average R^2: 0.40, range: 0.05 to 0.67). Average monthly sensible heat flux is 27% higher in NARR (range: -14% to 86%), and 2% lower in CLM outputs (range: -34% to 58%).

Monthly mean and standard deviation (inter-annual variability) of latent heat flux, sensible heat flux, and precipitation from AmeriFlux observations, NARR, and CLM outputs at the best available sites (longest comparison period, average 6 years, range: 5 – 7 years) in each region are shown in Figure 3.6. Seasonal variations (shape of monthly mean during the year) of latent heat flux and precipitation are captured by both NARR and CLM (see R^2 in Table 3.5). However, NARR latent heat flux shows higher positive bias compared to CLM as discussed previously (see bias in Table 3.5). Lower correlation of observed precipitation with model outputs, compared to temperature and latent heat flux, (Table 3.5) can be due to the multimodal (multiple peaks) nature of precipitation distribution during the year in Cfa, Dfa, and Dfb regions. The seasonal cycle of sensible heat flux show a bimodal pattern (two peaks during the year) at many AmeriFlux sites, particularly pronounced at the cropland site (e.g. US_Bo1_CRO in Fig. 3.6). CLM does not capture the bimodal pattern of sensible heat flux at all sites, and CLM shows only one peak in seasonal cycle of sensible heat flux. NARR captures the bimodal pattern of seasonal variations in sensible heat flux, particularly at US_Bo1_CRO site. The bimodal pattern issue is discussed in

detail in Section 3.5.4 and 3.5.5. NARR output and CLM input data show the observed seasonal cycle of near surface air temperature very well (not shown; see R^2 in Table 3.5). As shown in Fig. 3.6, regridding of NARR outputs did not change major characteristics of the results such as the bimodal pattern of sensible heat flux, higher positive bias in NARR latent heat flux compared to CLM, and multimodal nature of precipitation distribution at the four sites shown in Fig. 3.6.

Many studies have identified error in energy balance closure at FLUXNET sites (Wilson et al., 2002; Foken et al, 2008; Stockli et al, 2008). These errors are in the order of 20%, with underestimation of latent heat flux and sensible heat flux or overestimation of available energy [*Lht + Sht = 0.8 (Rn – Ght)*; Wilson et al., 2002]. Level 4 AmeriFlux data, that uses u* (friction velocity) filtering, is expected to show better energy balance closure, because improvements in energy balance closure are found with increasing friction velocity (Wilson et al., 2002). In the u* filtering method, measured fluxes below the threshold u* are discarded (mainly night time fluxes), and filled with the other observations with similar meteorological condition (gap filling). However, in Level 4 AmeriFlux data, net radiation or surface albedo variables are not available, and hence quantative evaluation of improvements in energy balance closure cannot be made at this time. In Section 3.5.6, NARR and CLM latent heat flux / *ET* outputs are evaluated using a theoretical approach (Budyko curve) instead of using AmeriFlux observations.

The results presented in this section are based on comparison of a point scale observation (25 x 25m) with NARR grid cell (32 x 32 km) and CLM grid cell (280 x 280 km) outputs. This comparison raises an important question about the validity of comparing point scale observations with coarse resolution gridded climate model output. This issue is addressed in Section.3.4.3.
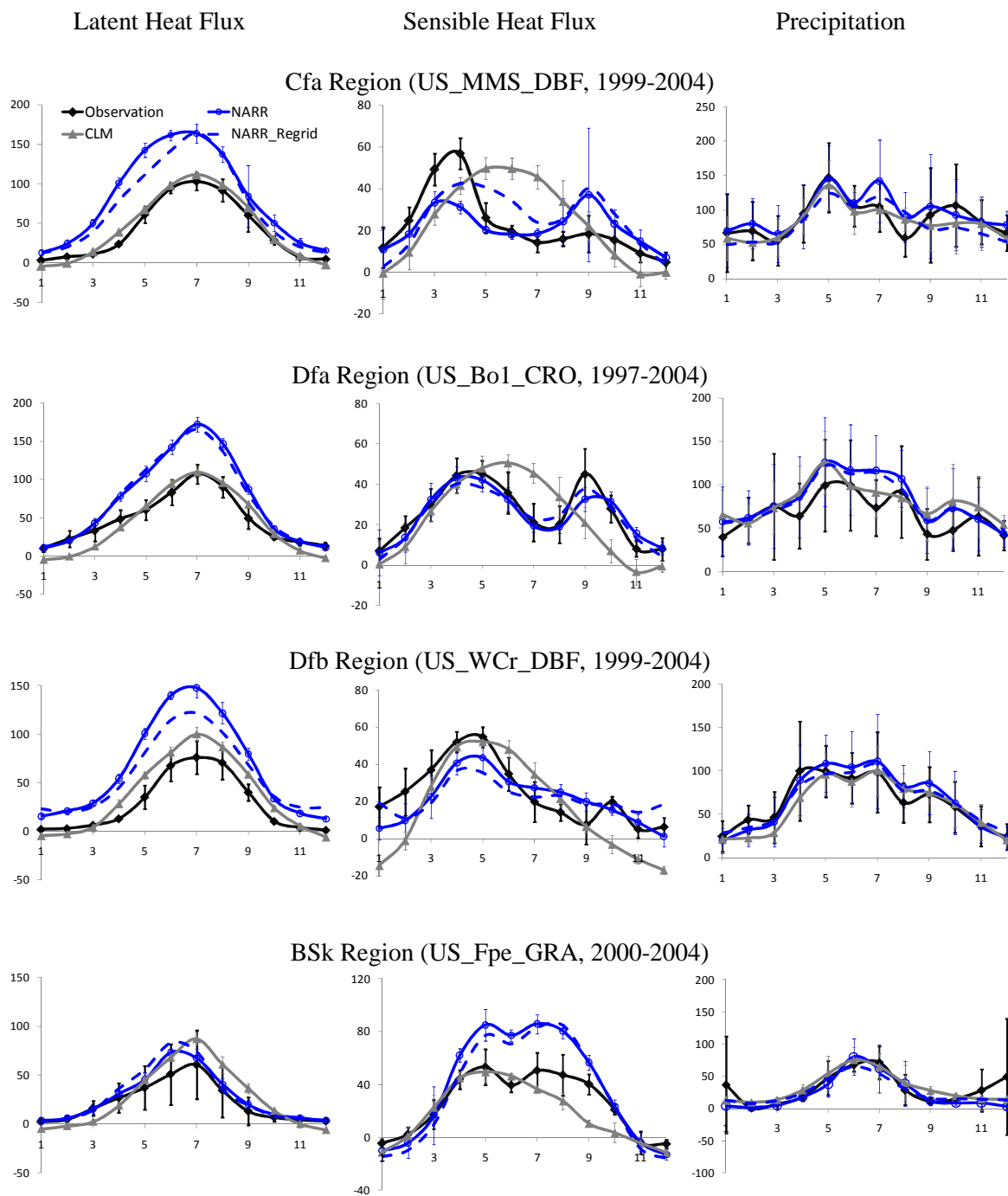
Fig. 3.6: Monthly climatology of latent and sensible heat fluxes and precipitation at four AmeriFlux sites, one each in Cfa, Dfa, Dfb, and BSk region. X-axis represent months 1 to 12 (Jan. to Dec.), and Y-axis represent monthly flux (W/m^2) for latent and sensible heat fluxes, and monthly precipitation (mm/month) for precipitation. Error bar represent $\pm$ 1 standard deviation in each case (AmeriFlux, CLM, and NARR). NARR_Regrid (blue broken line) is the NARR data regrided to T42 resolution as described in section 3.2. Only mean values of NARR_Regrid are shown.

3.4.3 Spatial Variability in Point Scale Hydroclimatic Observations

The issue of comparing point scale observation with climate model grid cell outputs is addressed by looking into the spatial variability of monthly precipitation and near surface air temperature records at 71 USHCN stations in Indiana and Illinois (Fig. 3.7). A total of 113 years (1896 to 2008) of monthly records are included in the analysis. Spatial variability is analyzed through pair-wise spatial correlation, RMSE difference, semi-variance, and statistical significance of difference in the monthly mean observation. The average distance among pairs of sites, and the number of station pairs for each distance is given in Table 3.6. The average distance ranges from 38 km to 524 km between any two stations, and a total of 2485 station pairs are included in the analysis.
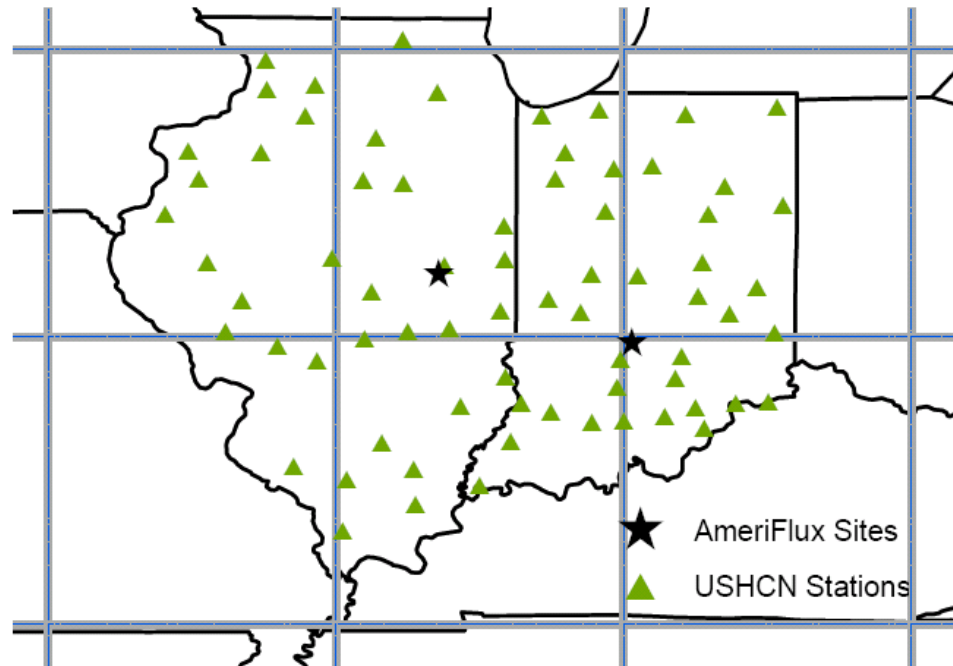
Fig. 3.7: Location of 71 USHCN stations in Indiana and Illinois with CLM grid (T42 resolution) in the background.

Table 3.6: Pairs of USHCN stations and their distances

Sl. No.	Distance (in km)	pairs of stations	mean distance (in km)
1	dist ≤ 50 km	58	38
2	50 km < dist ≤ 100 km	204	77
3	100 km < dist ≤ 200 km	635	153
4	200 km < dist ≤ 300 km	696	248
5	300 km < dist ≤ 400 km	552	345
6	400 km < dist ≤ 500 km	301	442
7	dist > 500 km	39	524

Spatial variability in monthly near surface air temperature records is shown in Fig. 3.8. Spatial correlation remains very high for all the distances (average correlation coefficient = 0.90), however its magnitude decreases with increasing distance (0.96 for 38 km distance and 0.82 for 524 km distance). The RMSE difference increases from 0.83 ^{O}C for 38 km distance to 3.2 ^{O}C for 524 km distance (average of 12 months). Semi-variogram of temperature data (Fig. 3.8 (c)) shows a parabolic model with no nugget effect, suggesting that the scale of variability is larger than the sampling interval (Kitanidis, 1997, p. 32-40). Statistical significance test (two sample T test) of difference in the means show that the monthly temperature records are not statistically different (α = 0.05) until an average distance of 248 km (Fig. 3.8 (d)). Spatial correlation and RMSE difference for 248 km distance are 0.91 and 1.8 ^{O}C, respectively.

Spatial variability in monthly precipitation records is shown in Fig. 3.9. Spatial correlation of precipitation records is lower than the spatial correlation of temperature records because of the higher variability and multimodal precipitation pattern in the region as discussed in Section 3.4.2. Spatial correlation of monthly precipitation decreases from 0.82 for 38 km distance to 0.30 for 524 km distance (average = 0.55). The RMSE difference increases from 29 mm/ month for 38 km distance to 64 mm/month for 524 km distance. The semi-variogram of monthly precipitation (Fig. 3.9 (c)) shows an almost linear shape with small nugget effect (410 mm/month ^2), suggesting that most of the variability is at a scale larger than the sampling interval, but some variability may be present at a scale comparable to the sampling interval (Kitanidis, 1997, p. 32-40). Nugget effect is a discontinuity of semi-variogram at the origin (y axis intercept), obtained by fitting a linear trend line to the mean semi-variogram curve shown in Fig. 3.9 (c). Statistical significance

test (two sample T test) shows that the mean precipitation is not statistically different for 7 months (April to Oct.) for all the distances. For five months (Nov. to March), mean precipitation becomes statistically different for 441 km or greater distances (Fig. 3.9 (d)). Spatial correlation and RMSE differences for 248 km distance are 0.53 and 50mm/month, respectively.
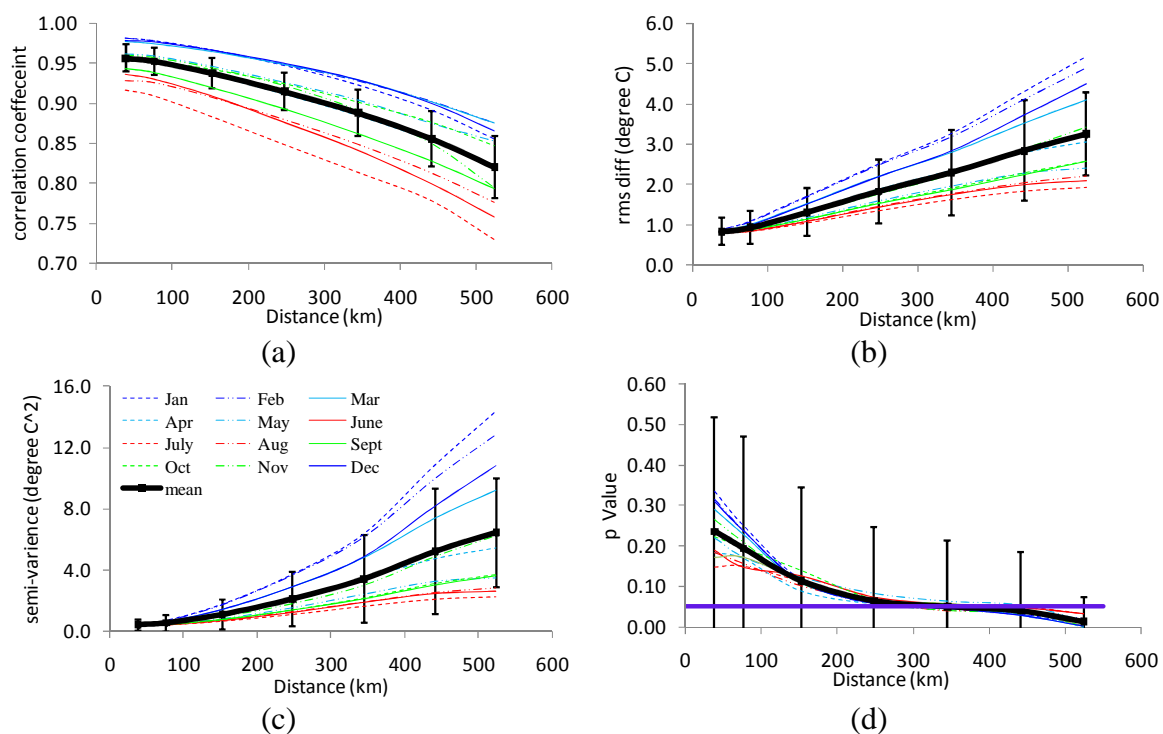
Fig. 3.8: Spatial variability in monthly temperature records at 71 USHCN stations in Indiana and Illinois. (a) spatial correlation, (b) root mean square difference, (c) semi-variance (d) statistical difference (p value). Thin lines represent mean value for each month, thick black line represent average value calculated from the monthly value, and error bars represent $\pm$ 1 standard deviation (average of 12 month).

The mean behavior of point scale observation of monthly precipitation and temperature records suggest that point scale measurements are not statistically different for at least 248 km distance in Indiana and Illinois region. Similar to precipitation and temperature observations, ubiquitous observations of surface energy fluxes are not available, and hence, spatial variability analysis as presented above cannot be conducted using sparse and short term energy flux observations. Pair-wise study is conducted using five or more years of comparative records from available neighboring stations to study the effects of climatic forcing, and land cover types on energy flux observations. The results from the pair-wise study are presented below, but it should be noted that

these results are obtained by using a small sample size (minimum sample size = 5 and maximum sample size = 9).
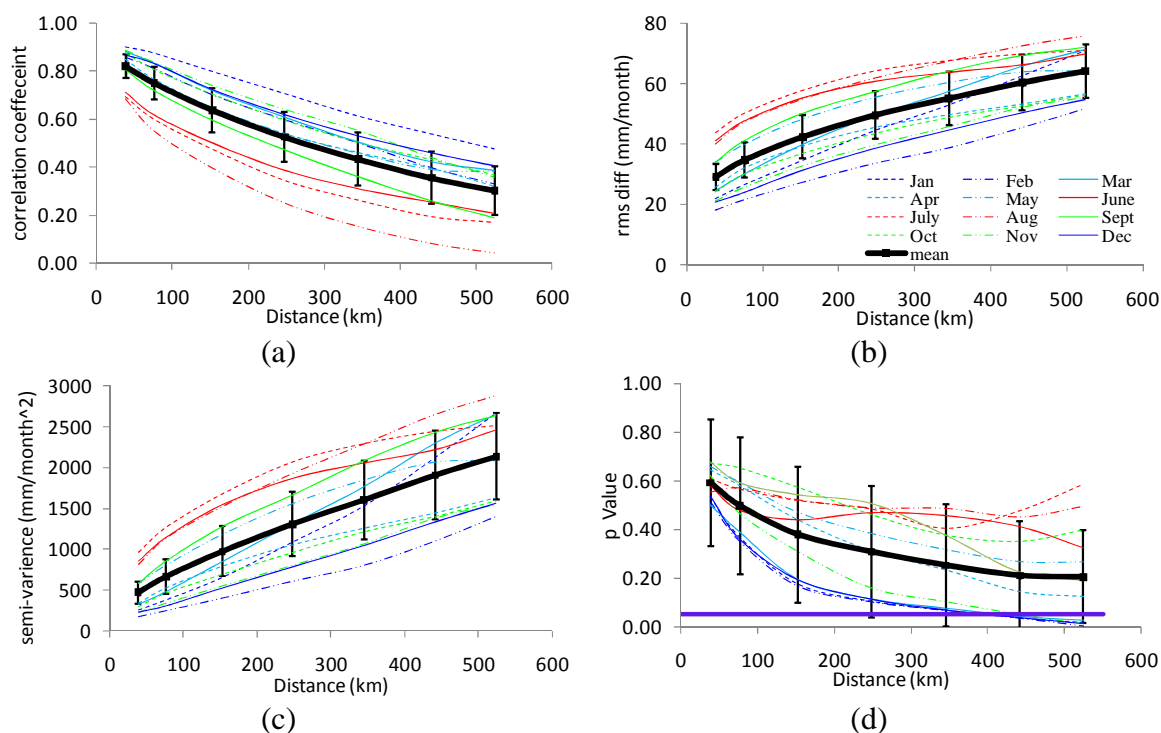
Fig. 3.9: Spatial variability in monthly precipitation records at 71 USHCN stations in Indiana and Illinois. (a) spatial correlation, (b) root mean square difference, (c) semi-variance (d) statistical difference (p value). Thin lines represent mean value for each month, thick black line represent average value calculated from the monthly value, and error bars represent $\pm$ 1 standard deviation (average of 12 month).

Monthly observations of latent and sensible heat fluxes along with temperature and precipitation are analyzed for three pairs of AmeriFlux sites including: (i) US_MMS_DBF, and US_BO1_CRO; (ii) US_Wcr_DBF, and US_Los_CSH; and (iii) US_Ha1_DBF and US_Ho1_ENF. The third pair of stations (US_Ha1_DBF & US_Ho1_ENF) located in the northeast United States is included in the analysis because of their longest available comparative records (9 years). Summarized results for summer (May to Oct.) and winter (Nov. to April) months are presented in Table 3.7.

AmeriFlux sites US_MMS_DBF and US_BO1_CRO are located 177 km apart. Average correlation coefficient for monthly temperature is 0.75 and 0.82 during summer and winter months, and is statistically significant for 4 months each in summer and winter. Correlation for

other variables is low compared to temperature correlation, and these correlations are not significant for most months. The summer months account for 87% of annual total *ET* for US_MMS_DBF, and 74% of annual total *ET* for US_BO1_CRO site. The magnitude of summer latent heat flux / *ET* is similar at both sites (73 W/m^2 at US_MMS_DBF site and 70 W/m^2 at US_BO1_CRO site), and they are not statistically different for any month in summer. Sensible heat flux is significantly different for most months in summer, and the average magnitude of sensible heat flux is 18 W/m^2 at US_MMS_DBF, and 32 W/m^2 at US_BO1_CRO. In winter, however, the behavior is opposite with statistically different latent heat flux for all months. Winter sensible heat flux is not statistically different for 4 months at US_MMS_DBF and US_BO1_CRO site.

AmeriFlux sites US_WCr_DBF and US_Los_CSH are located 32 km apart in north Wisconsin. Temperature data at these two sites show low correlation, 0.30 for summer, and 0.33 for winter. Low correlation in monthly temperature could be due to differences in elevation (Table 3.1) and terrain type (over-shaped ridge for US_Wcr_Dbf , and poorly drained depression/wetland for US_Los_CSH), but this issue is not investigated in this study. The summer months accounts for 90% of annual total *ET* at these sites, and latent heat flux / *ET* is of similar magnitude at both sites (48 W/m^2 at US_WCr_DBF, and 51 W/m^2 at US_Los_CSH site). Sensible heat flux is also similar in magnitude (27 W/m^2 at US_WCr_DBF, and 29 W/m^2 at US_Los_CSH), and is not statistically different for 5 months in summer. During winter, sensible heat flux has different magnitude at these sites (26 W/m^2 at US_WCr_DBF, and 17 W/m^2 at US_Los_CSH), but the difference is not statistically significant for four months.

AmeriFlux sites US_Ha1_DBF (in Massachusetts), and US_Ho1_ENF (in Maine) are located 405 km apart. Monthly temperature shows correlation of 0.72 in summer and 0.84 in winter. Other variables show relatively lower correlation (e.g. 0.38 for latent heat flux during the summer months). The summer months account for more than 80% of annual total *ET* at these sites. Summer latent heat flux / *ET* at US_Ha1_DBF is 56 W/m^2, and 49 W/m^2 at US_Ho1_ENF site (significantly different for two months). The sensible heat flux at US_Ha1_DBF and US_Ho1_ENF sites is statistically different for most months in summer, and for three months in winter.

Table 3.7: Pair wise comparative analysis of AmeriFlux observations at selected sites (T_{air} /P/Lht/Sht; units are ^{O}C/mm per month/W per m^2/W per m^2).

(a) For summer months (May to October, 6 months)

Sl. No.	Group	Distance (KM)	Comparison year	Correlation Coefficient	NSSCM	Mean Value	NSDMM
1	US_MMS_DBF	177	1999-2006*	0.75/0.23/-0.07/0.32	4/1/0/0	19.9/100/73/18	1/1/0/4
	US_BO1_CRO					19.3/77/70/32	
2	US_WCr_DBF	32	2001-2005	0.30/NA/0.21/0.40	0/NA/1/0	14.3/NA/48/27	0/NA/1/1
	US_Los_CSH					13.7/NA/51/29	
3	US_Ha1_DBF	405	1996-2004	0.72/0.57/0.38/0.19	5/3/0/0	15.6/94/56/34	1/0/2/5
	US_Ho1_ENF					15.1/79/49/44	

Table 3.7: Pair wise comparative analysis of AmeriFlux observations at selected sites (T_{air} /P/Lht/Sht; units are ^{O}C/mm per month/W per m^{2}/W per m^{2}; continued)

(b) For winter months (Nov. to April, 6 months)

Sl. No.	Group	Distance (KM)	Comparison year	Correlation Coefficient	NSSCM	Mean Value	NSDMM
1	US_MMS_DBF	177	1999-2006*	0.82/0.57/0.53/0.46	4/1/1/1	4.9/79/10/26	0/0/6/2
	US_BO1_CRO					3.6/61/25/20	
2	US_WCr_DBF	32	2001-2005	0.33/NA/0.25/0.50	0/NA/1/2	-3.5/NA/4/26	0/NA/0/2
	US_Los_CSH					-4.5/NA/6/17	
3	US_Ha1_DBF	405	1996-2004	0.84/0.39/0.38/0.24	6/2/2/0	0.2/87/12/30	5/2/3/3
	US_Ho1_ENF					-2.0/57/8/19	

Values are calculated for each 12 months separately, and then averaged/counted for summer and winter months, respectively. NSSCM: Number of statistically significant correlation month (p value $\leq$ 0.05), NSSDMM: Number of significantly (statistical) different mean month (p value $\leq$ 0.05). * year 2000 is missing.

3.4.4   Spatial and Temporal Distribution of Total Runoff in MRB

Spatial distribution of annual total runoff based on UNH-GRDC composite runoff data, VIC model, NARR, and CLM is presented in Fig. 3.10. Except for the UNH-GRDC dataset, for which only monthly climatological mean values are available (Fekete et al., 2002), the other three model outputs represent average annual total runoff from 1988 to 1999. Data from 1988 to 1999 are presented because VIC outputs are available for that period only. All runoff datasets are shown at their original resolutions, i.e., UNH-GRDC at 0.5^0 resolutions (~ 50 km), VIC at 0.25^0 resolution (~ 25 km), NARR at 32 km resolution, and CLM at T42 resolution (~ 280 km). The UNH-GRDC data shows some discontinuity (lower runoff) in western Kentucky. The VIC output follows the precipitation gradient shown in Fig. 3.2(b), and provides better geographic distribution of runoff in MRB (Roads et al., 2003). The spatial distribution of total runoff from CLM is visually comparable to that from VIC; whereas NARR provides lower total runoff in all area of MRB.

The intra-annual variability of total runoff is shown in Fig. 3.11 (a). Naturalized observed runoff from Maurer and Lettenmaier (2001) is also included in the analysis. The monthly runoff from VIC is closer to the observation, with an RMSE of 2.7 mm/month. The UNH-GRDC data show lower monthly runoff from June to December, and the annual average RMSE for UNH-GRDC data is 6.2 mm/month. The monthly runoff from CLM is higher during winter months, and is comparable with VIC runoff during summer months. Compared to the observed runoff data from Maurer and Lettenmaier (2001), annual average RMSE in CLM monthly runoff is 7.7 mm/month. The runoff data from NARR is lower for all months, and the annual average RMSE in NARR is 12.7 mm/month compared to Maurer and Lettenmaier (2001).

The annual time series of total runoff is shown in Fig. 3.11 (b), and the annual total runoff statistics are presented in Table 3.8. VIC output data match closely with observations, with a correlation coefficient of 0.94, bias of 2%, and RMSE of 13 mm/year. The total runoff from CLM is higher for all years, with a bias of 19%, and RMSE of 47 mm/year. Higher annual runoff from CLM may be due to higher runoff during winter months as mentioned in the previous paragraph. Overall, CLM total runoff captures the inter-annual variability, with a correlation coefficient of 0.91. CLM total runoff results found in this study are consistent with findings of Oleson et al. (2008), who also found higher overestimation of total runoff and high correlation coefficient. NARR gives lower total runoff for all years, with a bias of -62% and RMSE of 151 mm/year. In addition, NARR data show relatively poorer correlation (correlation coefficient: 0.52) in

comparison with CLM. Regridding of NARR data did not affect monthly or annual runoff results in MRB (Fig. 3.11).
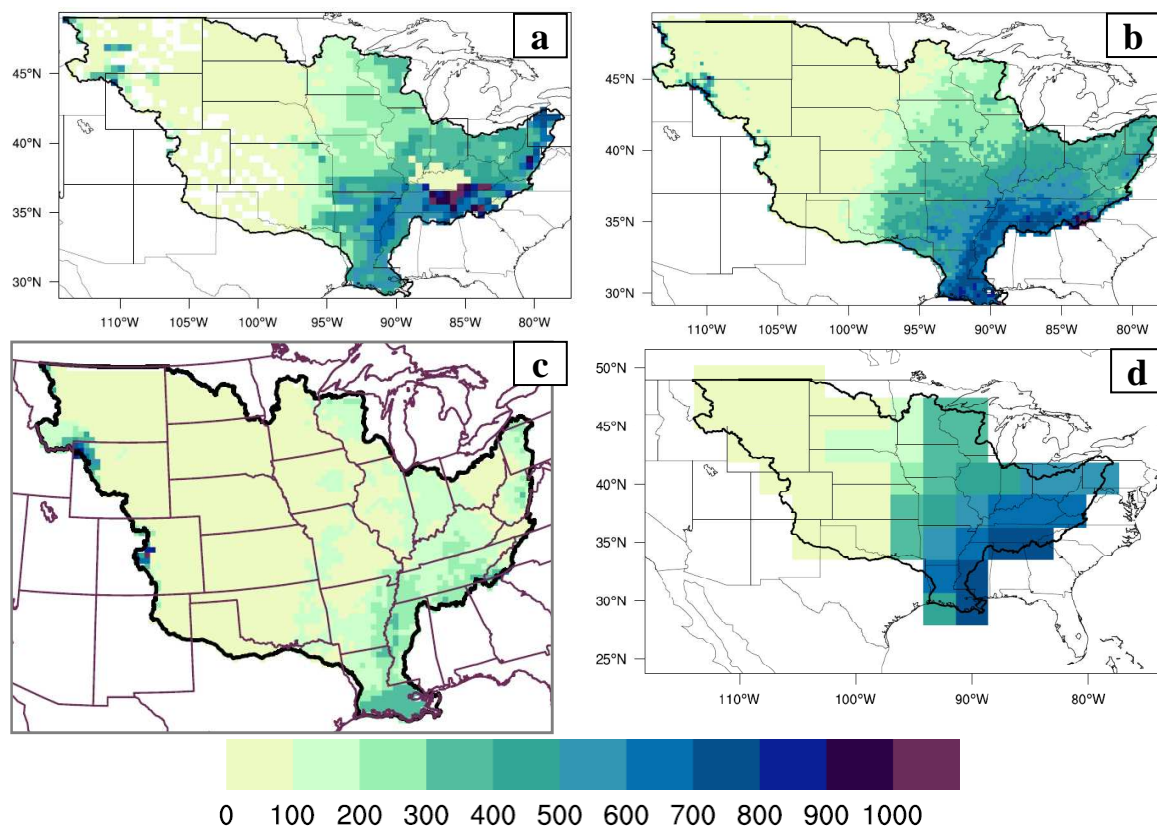
Fig. 3.10: Spatial distribution of annual total runoff (mm/year) in MRB (a) UNH-GRDC composite runoff data (b) VIC model [1988-1999] (c) NARR [1988-1999] (d) CLM [1988-1999].

Table 3.8: Annual total runoff statistics (1988 – 1999) in MRB

ID	mean mm/year	Correlation coefficient	RMSE mm/year
Observation	237		
UNH-GRDC	187		
VIC	242	0.94	13
NARR	89	0.57	151
CLM3.5	281	0.91	47
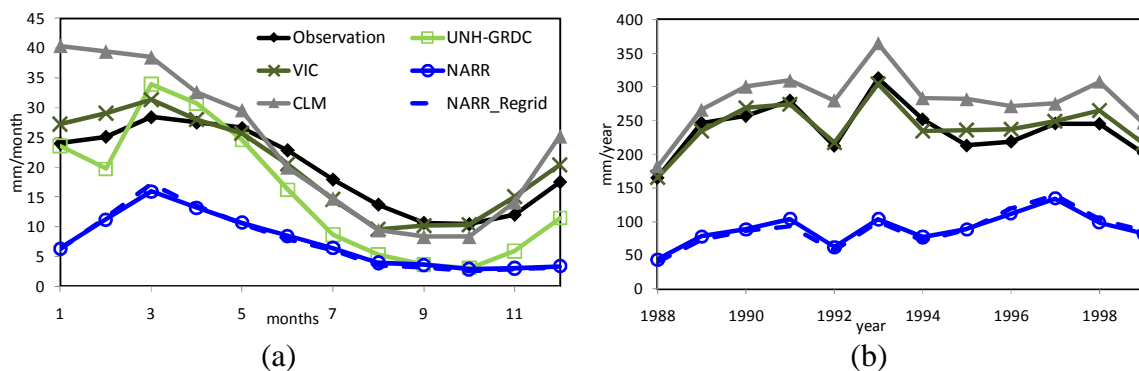
Fig. 3.11: Temporal distribution of total runoff in MRB (a) Intra-annual variability (b) inter-annual variability. NARR_Regrid (blue broken line) is the NARR data regrided to T42 resolution as described in section 3.2. Observation is from Maurer and Lettenmaier (2001).

3.4.5   Results Summary

NARR and CLM outputs are evaluated for surface water and energy fluxes in MRB using energy flux observations, and other relevant data/model outputs (e.g. runoff observations). Monthly climatology of near surface air temperature and precipitation of NARR is comparable to CLM and PRISM data. However, sensible heat flux and latent heat flux differ significantly between NARR and CLM outputs. Compared to average AmeriFlux data from 11 sites, NARR shows relatively higher biases in incoming solar radiation (24%), sensible heat flux (27%), and latent heat flux (59%); whereas CLM shows relatively smaller biases in incoming solar radiation (0.5%), sensible heat flux (-2%), and latent heat flux (11%). Similarly, annual and monthly total runoff is also better simulated by CLM compared to NARR. Based on 25 years (1980 – 2004) monthly climatology water and energy balance components in MRB, NARR has 11% energy balance closing error ($Lht + Sht + Ght = 1.11\ R_n$) and 12% water balance closing error ($ET + N = 1.12\ P$); whereas CLM does not have water and energy balance closing error by virtue of model design. Net radiation in NARR (84.7 W/m^2) is higher compared to CLM (69.7 W/m^2). Overall, CLM outputs provide better characterization of surface water and energy fluxes in MRB.

The issue of comparing point scale observations with gridded model outputs is addressed by using 113 years of monthly precipitation and temperature records at 71 USHCN stations in Indiana and Illinois. It is found that monthly precipitation and temperature are not statistically different for at least 248 km distance in Indiana and Illinois. Analysis of pair-wise energy flux

observations from neighboring stations show that effects of land cover type on summer latent heat flux/*ET* (which is greater than 80% of annual total *ET*) is minimal. Sensible heat flux show higher difference compared to latent heat flux at neighboring stations with different land cover types.

3.5 Discussion

3.5.1 Reanalysis as Surrogate for Observations

Reanalysis data / outputs are often used as a surrogate for observations for verification of climate model outputs (Covey et al., 2003; Gates et al., 1999; Kumar et al, 2010; Lambert and Boer, 2001; Richler and Kim,2008; http://www.cgd.ucar.edu/cms/diagnostics/), and other relevant hydroclimatic studies (Dominguez and Kumar, 2008; Dominguez et al, 2008; Diffenbaugh, 2009; Fall et al., 2009). This study used an improved version regional reanalysis product (NARR, Mesinger et al., 2006), and found that while precipitation and near surface air temperature are comparable to observed data, *ET* and runoff outputs have significant biases. The water and energy balance error observed in NARR could be due to: (i) assimilation of a limited number of available hydroclimatic variables in the NARR system, (ii) biases in net radiation, and (iii) parameterization of surface energy fluxes in NARR. While precipitation is an assimilated variable in NARR, sensible heat flux, latent heat flux and runoff are not assimilated in NARR. This finding suggests that reanalysis fields for which observations are not assimilated (e.g. sensible and latent heat flux) should be used with caution. Some aspects of surface energy flux parameterization difference between NARR and CLM are discussed in Section 3.5.5.

3.5.2 Point Scale Observations versus Gridded Model Output

Reanalysis data or other climate model outputs have the advantage of being continuous in spatial and temporal domain. In contrast, AmeriFlux data are sparse in spatial domain, and are available for a relatively short temporal domain (less than 15 years). This study show that CLM outputs are comparable to point scale energy flux observations at many sites, and CLM also produces better results (e.g. runoff for the MRB) at the continental basin scale (3.2 million sq. km) compared to

NARR. Hence, in addition to providing valuable information for model development (Stockli et al, 2008), point scale observations are valuable dataset for evaluation or verification of global and regional climate model outputs. Point scale energy flux observations cannot replace reanalysis outputs, but can provide first order assessment for both reanalysis data and climate model outputs. Hence, point scale observations should be taken into consideration alongside available reanalysis outputs in hydroclimatic studies.

Comparison of point scale observations with climate model grid cell outputs also brings up the issue of spatial scale, and how the heterogeneity in topography, land cover and soil within a model grid cell are captured in a point measurement. Results from this study show that the issue of scale and heterogeneity are masked at monthly time scale over the grid cell size of 280 km used in this study. Vinnikov et al. (1996) have proposed statistical models for spatial and temporal variability for soil moisture observations in the mid-latitude region (former USSR), having similar formulation (first-order Markov process). Temporal averaging within a month in the same year, and averaging during the same months for different years could be compensating for the spatial variability within a grid cell for *ET* or latent heat flux measurements shown in Fig. 3.6. This, however, may change for different variables in different regions such as mountainous regions.

This study also found that the correspondence between point scale measurement and CLM grid cell output is poor in the Rocky mountain range (US_NR1_ENF site). Some past studies (e.g., Han and Roads, 2004) have shown that a high resolution climate model performs better in a mountainous region compared to a low resolution climate model such as T42 resolution (280 km) in CLM. The spatial variability analysis conducted in this study using precipitation and temperature data from Indiana and Illinois (relatively flat topography) may not remain valid for the highly variable topographic regions (e.g. Rocky mountain region). Higher resolution climatic models may be needed for similar analysis in mountainous regions. For example, NARR (32 km resolution) performance is better for precipitation and temperature compared to CLM forcing data (280 km resolution) at US_NR1_ENF site (Table 3.5).

3.5.3   Effect of Land Cover Type on Sensible and Latent Heat Fluxes

Point scale measurements of sensible heat flux are found to be less correlated with the nearest climate model grid cell outputs compared to latent heat flux measurements (e.g. average correlation coefficient of CLM and AmeriFlux is 0.80 for latent heat flux, and 0.40 for sensible heat flux; also see Randerson et al., 2009). Pair-wise comparison of AmeriFlux observations (Table 3.7) show that sensible heat flux has a higher difference compared to latent heat flux at neighboring stations having different land cover types (e.g. summer average latent heat flux and sensible heat flux for US_MMS_DBF/ US_BO1_CRO sites are 73/70 and 18/32 W/m^2, respectively). These results suggest that land surface heterogeneity (e.g. land cover type) has higher effects on sensible heat flux compared to latent heat flux; i.e. land surface hydrologic response (*ET*/latent heat flux) is more stable compared to the land surface thermal response (sensible heat flux). This finding is consistent with some previous studies including the African Monsoon Multidisciplinary Analysis (AMMA) project where Ramier et al. (2009) found that the stability in *ET* is achieved at the expense of variability in other energy (e.g., sensible heat flux) and water balance components (e.g., soil water storage). Similarly, by using six years of energy flux data in a Mediterranean climate region (central California), Ryu et al. (2008) found that the inter-annual variability in ET is much less compared to the two fold change in annual precipitation during the observation period.

3.5.4   Bimodal Pattern of Sensible Heat Flux

A bimodal pattern in monthly climatology of sensible heat flux is found at four crop cover AmeriFlux sites in Dfa region. For one crop cover AmeriFlux sites in Cfa region (US_ARM_CRO), bimodal pattern of sensible heat flux was not found. For two irrigated crop cover sites (US_Ne1_CRO, and US_Ne2_CRO), mean value of sensible heat flux is nearly zero or slightly negative during July and August months, which also coincides with highest latent heat flux months in the year (not shown). This special phenomenon was explained by Tanner and Lemon (1962) and Monteith (1965) as the effect of advection of dry air over cooler and water sufficient crop area. *ET* from irrigated crop area exceeds net radiation, and hence temperature of irrigated crop area is lower than the air temperature. The deficit in the energy demand is met by the downwind transfer of sensible heat flux from the warmer air into the cooler crop area

(negative sensible heat flux). During the senescence period, *ET* decreases, and gradually sensible heat flux becomes the dominant energy flux (Li et al, 2005).

3.5.5   Model Parametrization Differences NARR and CLM

NARR captures bimodal pattern of sensible heat flux at crop sites (e.g. US_Bol_CRO), whereas CLM does not capture the bimodal pattern of sensible heat flux (Fig. 6b). NARR uses the Noah land surface parameterization scheme (Chen and Dudhia, 2001), where sensible heat flux is calculated as the residual of the energy balance terms ($Sht = R_n – Lht – Ght$; Sridhar et al., 2002), hence *Sht* can become uncoupled from temperature seasonality (one peak during the year) as *Lht + Ght* approaches to *Rn*. Whereas in CLM parameterization, sensible heat flux is calculated as heat transfer from ground/vegetated surface to atmosphere using the temperature difference between the atmosphere and ground/vegetated surface as the potential difference and aerodynamic resistance as resistance based on Ohm's law (Oleson et al., 2004, page 56-64). Hence, it is less likely that *Sht* can become uncoupled from temperature seasonality in CLM. Ground heat flux is calculated as the residual of the energy balance terms in CLM ($Ght = R_n – Lht – Sht$; Oleson et al., 2004, page 77). Although the opposing signs of ground heat flux cancel each other at longer time scale (e.g. annual), ground heat flux can be substantial at daily time scale. Ramier et al. (2009) show the mean absolute value of ground heat flux as less than 7% of (*Lht + Ght*) based on two years of observations in the AMMA project. During wet spells, soil releases the heat to the atmosphere; whereas during the dry spell, soil gains heat from the atmosphere (Ramier et al., 2009).

Based on a point scale model run at three flux tower sites, Wang et al. (2008) found that the ground heat flux is poorly simulated compared to latent heat flux and sensible heat flux in the CLM (version 3) model. On the other hand, Sridhar et al. (2002) found good correspondence (statistically significant strong correlation) between measured and modeled ground heat flux at 7 flux measurement sites in Oklahoma using the Noah land surface model. There could be other parameterization differences that may be contributing to the differences in sensible heat flux monthly climatology that is seen in this study. Further investigation of this issue may require running the CLM and Noah land surface model at point scale. Santanello et al. (2009) have compared CLM and Noah land surface model coupled with different Planetary Boundary Layer

schemes, but their analysis results are limited to diurnal time scale, and are not sufficient to analyze the seasonal cycle.

3.5.6   Evaluation of NARR and CLM ET using Budyko Curve

Large differences between NARR and CLM latent heat fluxes in the eastern part of MRB (Fig. 3.4), and energy balance closure issue with AmeriFlux observations (Section 3.4.2) warrant further investigation of this issue using some independent approach. Budyko Curve is a top-down approach for basin average ET estimation as opposed to bottom-up process based approach e.g. NARR and CLM. Budyko postulated that for long term average, under very dry conditions actual ET is limited by precipitation, and under very wet conditions actual ET is limited by available energy (Budyko 1958). Between these two limits a number of curves have been proposed to account for increasing complexity e.g. catchment characteristics and finer temporal scale (Milly, 1994; Koster and Saurez 1999; Zhang et al, 2001, 2004, and 2008). Fig. 3.12 shows Budyko curve given in Eq. 3.5. This curve is applicable for the steady state condition ($\Delta s = 0$ in Eq. 3.1, long term annual mean), and its validation over 250 catchments has been shown in Zhang et al. (2001).

$$\frac{ET}{P} = \frac{1 + w \cdot \Phi}{1 + w \cdot \Phi + \Phi^{-1}} \tag{3.5}$$

where Φ $(= \dfrac{R_n}{\lambda \cdot P})$ is the dryness index, and w is the plant available water coefficient ($w = 2.0$ for forest, and $w = 0.5$ for short grass and crops).

The performance of CLM and NARR ET in the Ohio and Tennessee Basin (the eastern part of the MRB, Region 05 and 06 Fig. 2(c)) are evaluated for long term annual mean (1980 – 2004; Fig. 3.12). The observed mean annual ET for the Ohio and Tennessee basins is calculated as the difference between observed mean annual precipitation (using PRISM data) and the observed mean annual streamflow data at United States Geological Survey (USGS) gauging station # 03611500 (Ohio river at Metropolis, Illinois with a drainage area of 0.52 million sq. km). The observation falls on the Budyko curve ($ET/P = 0.58$, and $\Phi = 0.80$), whereas NARR shows an overestimation of ET ($ET/P = 0.92$, and $\Phi = 1.12$), and CLM show underestimation of ET ($ET/P = 0.35$, and $\Phi = 0.85$). Observed ET estimates can have some positive bias due to flow

regulations upstream of the gauging station (~ 5%, Maurer and Lettenmaier, 2001; USGS Water Data Report 2009). Lower *ET/P* in CLM is also consistent with the overestimation of CLM runoff as discussed in Section 3.4.4. Budyko approach also confirms that NARR overestimates ET in MRB.
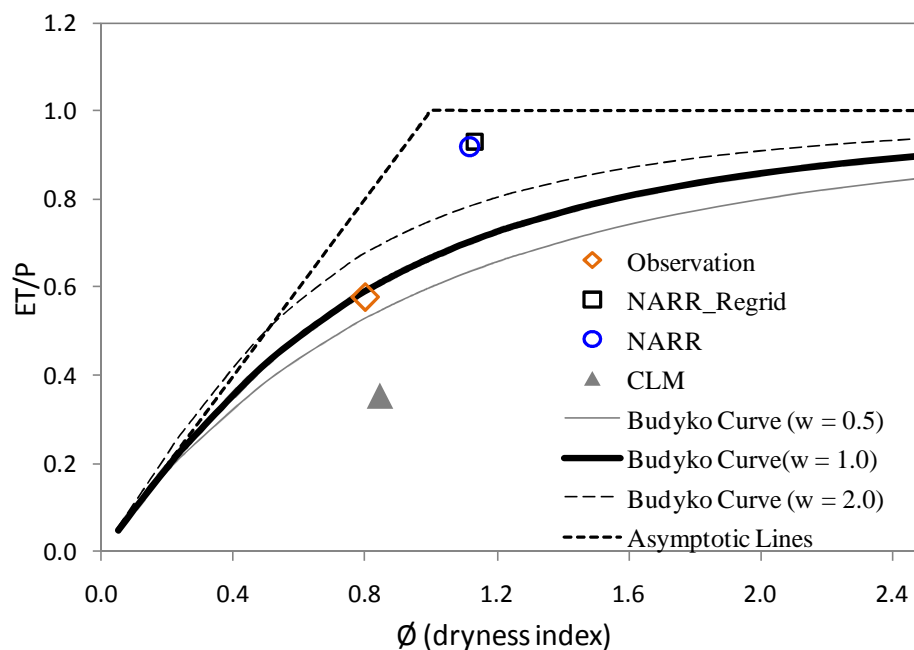
Fig. 3.12: Performance evaluation NARR and CLM ET outputs using Budyko curve in Ohio-Tennessee basin for 1980-2004 annual average.

3.5.7   Comparison with the WEBS study

A comparison of this study with the WEBS study for major surface water and energy budget terms (ET/P, N/P, Lht/Rn, Sht/Rn) is presented in Table 3.9. Partitioning of precipitation into ET and runoff has not significantly improved in NARR as compared to REAN1 and REAN2. ET approximately balances total precipitation in all three reanalysis products. In NARR, water balance closing error (ET + N = 1.12 P) is similar to REAN2 (ET + N = 1.11P). It seems that runoff is not an important variable for climate models, as also seen in GSM, RSM and ETA results (The WEBS study). CLM which is run in offline mode produces results comparable to the VIC model. CLM coupled model run (coupled with atmosphere and ocean components) is not available for evaluation. Latent heat flux has decreased and sensible heat flux has increased in

NARR compared to REAN1, REAN2, GSM, and RSM. The issue of negative sensible heat flux during winter as identified in the WEBS study has improved in NARR (Fig. 3.5 and 3.6).

Table 3.9: Comparison with the WEBS study.

	The WEBS Study						This Study		
	REAN1	**REAN2**	**GSM**	**RSM**	**ETA**	**VIC**	**NARR**	**CLM**	**OBS**
ET/P	1.02	1.05	0.89	0.88	0.94	0.71	1.00	0.66	0.71
N/P		0.06	0.15	0.09	0.17	0.29	0.12	0.34	0.29
Lht/Rn	0.77	0.85	0.75	0.75	0.64	0.63	0.70	0.59	
Sht/Rn	0.13	0.10	0.17	0.28	0.34	0.33	0.40	0.39	

The WEBS climatology is for 1996 to 1999, NARR and CLM (This study) is for 1980 – 2004, and the observation (OBS) is for 1988 to 1999. Observed ET is estimated as the difference between average PRISM precipitation and average naturalized runoff for 1988 to 1999. Details of the WEBS study are given in Roads et al. (2003).

3.6 Concluding Remarks

This study highlights the differences in climate model output (CLM), observations (AmeriFlux) and reanalysis products (NARR) in hydroclimatic assessments at continental basin scale. Hydroclimatic variables for which observations are not assimilated in the reanalysis products (e.g. *ET* and runoff in NARR) should be used with caution for evaluation of climate model outputs. For example, evaluation of CLM using NARR only may show that CLM underestimates *ET* or latent heat flux, and overestimates sensible heat flux in MRB (Fig. 3.5). However, this is not the case as found in this study because CLM ET output matches more closely with AmeriFlux data compared to NARR *ET* (Fig. 3.6).

The availability of AmeriFlux observations in recent years has proved to be an important data source to improve our understanding of land surface and atmospheric interaction such as variability in latent heat flux compared to sensible heat flux and bimodal seasonal cycle of sensible heat flux as found in this study for MRB. Issues related to scaling (point scale measurement versus climate model grid cell outputs) and energy balance closure in Level 4 AmeriFlux data should be addressed in future studies so that AmeriFlux observations can be used more extensively in hydroclimatic studies.

Spatial variability analysis of monthly precipitation and temperature using 71 USHCN in Indiana and Illinois show that monthly precipitation and temperature vary at a scale larger than the average distance (39 km) between any two neighboring stations in this study. Gridded network of AmeriFlux sites is needed to conduct similar studies for latent and sensible heat fluxes. Pair-wise analysis of energy flux observations at neighboring AmeriFlux sites in this study is limited by small sample size. Field experiments similar to the AMMA project (same climatic condition and different land cover type, Ramier et al., 2009) in the mid-latitude region can give more information on effects of land surface heterogeneity on sensible versus latent heat fluxes.

Surface energy flux parametrization and formulation differences between the NOAH land surface model (land component of NARR) and CLM discussed in this paper indicates room for further improvement in CLM (how to capture bimodal seasonal cycle of sensible heat flux, underestimation of ET in the eastern domain). Basin scale energy and water balance study as well as the Budyko approach show overestimation of *ET* in NARR.

NARR data show significantly lower (62%) total runoff in MRB. This finding is consistent with other studies related to hydrologic validation of reanalysis data (Hagemann et al., 2005; Lucarini et al., 2007), as well as the WEBS study. Assimilation of observed runoff data may address this issue in future reanalysis projects. CLM has relatively improved runoff output that is comparable to a mesoscale hydrologic model runoff output (VIC in this study).

This study is constrained by the limited number of models (CLM and NARR) and only one study region (MRB). Similar studies incorporating other models as well as other study areas will help to bring more confidence in the global climate model simulation results, which can then be used for regional scale water resources planning and management.

3.7 <u>Summary</u>

The North American Regional Reanalysis (NARR) and Community Land Model (CLM, version 3.5) outputs are analyzed to characterize the surface water and energy budgets in the Mississippi River Basin (MRB). NARR and CLM performance are evaluated with reference to energy flux observations from 16 AmeriFlux sites in MRB. The issue of point scale observations versus climate model grid cell outputs is addressed by analyzing the spatial variability in long term monthly precipitation and temperature observations from 71 United States Historical Climatology Network stations in Indiana and Illinois. The model outputs are also evaluated for their ability to capture spatial and temporal variability in total runoff. Compared to average values at 11 AmeriFlux sites in MRB, NARR show higher biases (compared to CLM) in incoming solar radiation (24%), sensible heat flux (27%), and latent heat flux (59%); whereas, CLM show smaller biases (compared to NARR) in incoming solar radiation (0.5%), sensible heat flux (-2%), and latent heat flux (11%). Seasonal cycle of observed sensible heat flux in the crop region show two peaks (bimodal pattern), which is captured by NARR, but CLM does not show any bimodal pattern. Based on 25 years (1980 – 2004) monthly climatology in MRB, NARR has 11% energy balance closing error (*latent + sensible + ground heat flux = 1.11 net radiation*) and 12% water balance closing error (*evapotranspiration + runoff = 1.12 precipitation*); whereas CLM does not have water and energy balance closing errors, primarily due to model design. In comparison to the observed mean annual runoff of 237 mm/yr based on 1988-1999 data in MRB, NARR and CLM mean annual runoff values are 89 mm/year and 281 mm/year, respectively. Overall, CLM provides relatively better characterization of surface water and energy fluxes in the MRB compared to NARR.

CHAPTER4. LAND COVER CHANGE IN THE UNITED STATES

4.1 <u>Introduction</u>

Land-cover change is an important phenomenon that has occurred globally over the past several centuries. Between 1700 and 2000, the global extent of natural vegetation has decreased by 45%, and agricultural land area has increased by 500% (Pongratz et al., 2008, Scanlon et al., 2007). Land-cover change has several important implications including: (i) changes in global and regional climate (Bonan, 1997, 1999; Brovkin et al, 2004; Kumar et al., 2010; Lawrence and Chase, 2010; Matthews et al., 2004; Pielke et al., 2002; Pielke and Niyogi, 2010); (ii) changes in the hydrologic cycle through evapotranspiration, runoff, and irrigation (Baldock et al., 2000; Gordon et al., 2005; Yang et al., 2010; Zhang and Schilling, 2006); (iii) acceleration of the biogeochemistry (Gruber and Galloway, 2008; Basu et al., 2010); and (iv) fragmentation and/or loss of habitats (Carlisle et al., 2010). While a number of studies have documented historical land-cover change (Goldewijk and Drecht, 2006; Pongratz et al., 2008; Ramankutty and Foley 1999), the underlying governing mechanisms of land-cover change has not been explored in sufficient detail. For example, how the contribution of different driving forces (e.g. socio-economic versus biophysical drivers) have changed through time remains to be investigated.

The environmental change history of the United States is relatively new (1600 to present) and well documented. For example, high resolution (county level) population data since 1790 are available, and   agricultural data since 1850 are available (Pastore et al., 2010; Waisanen and Bliss, 2002; Whiteny, 1994; Williams, 1989). A large geographic area of the conterminous United States (8.08 million km^2) with variable topography and climate (84 level-III ecoregions; Omernik, 1987), and settlement history during the 19th and 20th century offer an opportunity to explore contributions of land-cover change drivers in the United States. Based on the knowledge of the agricultural history of the United States (Fig. 4.1, and for details see Section 4.2), this study attempts to relate the historical spatial and temporal changes in cropland cover from 1850 to 2000 to socio-economic and biophysical drivers. More simply, we have attempted to answer the

question – why do we see what we see (Fig. 4.1)? Understanding contributions of land-cover change drivers will be helpful for the modeling of future land-cover change scenarios.

Studies related to the global historical distribution of cropland (1700 to present) are mainly based on coarse-resolution crop-inventory data (national/sub-national scale, Ramankutty and Foley, 1999), and population density as the primary driver for cropland distribution (Historical Database for the Global Environment [HYDE3], Goldewijk and Drecht, 2006). HYDE3 study found the population density to be an unsuitable proxy for cropland distribution during the 20th century. Unlike these coarse-resolution global-scale studies, the USGS (United States Geological Survey) Land Cover Trends Project is very high resolution and data intensive (based on 10 km x 10 km random samples in each ecoregion, and change detection from Landsat images; Loveland et al., 1999). However, the Trends project provides information from 1973 – 2000 only, thus making it inadequate to answer the question posed in this study.

The specific objectives of this study, which is primarily designed on the reverse engineering concept (object to theory), are to: (i) quantify the contribution of drivers for cropland spatial distribution in the conterminous United States from 1850 to 2000; (ii) study cropland trajectory and its variations among ecoregions by fitting a hybrid analytical function to normalized cropland % (NCP, see Eq.4.1) in each ecoregion; and (iii) study the effect of inter-annual, and inter-decadal climate variability (precipitation) on the cropland trajectory.
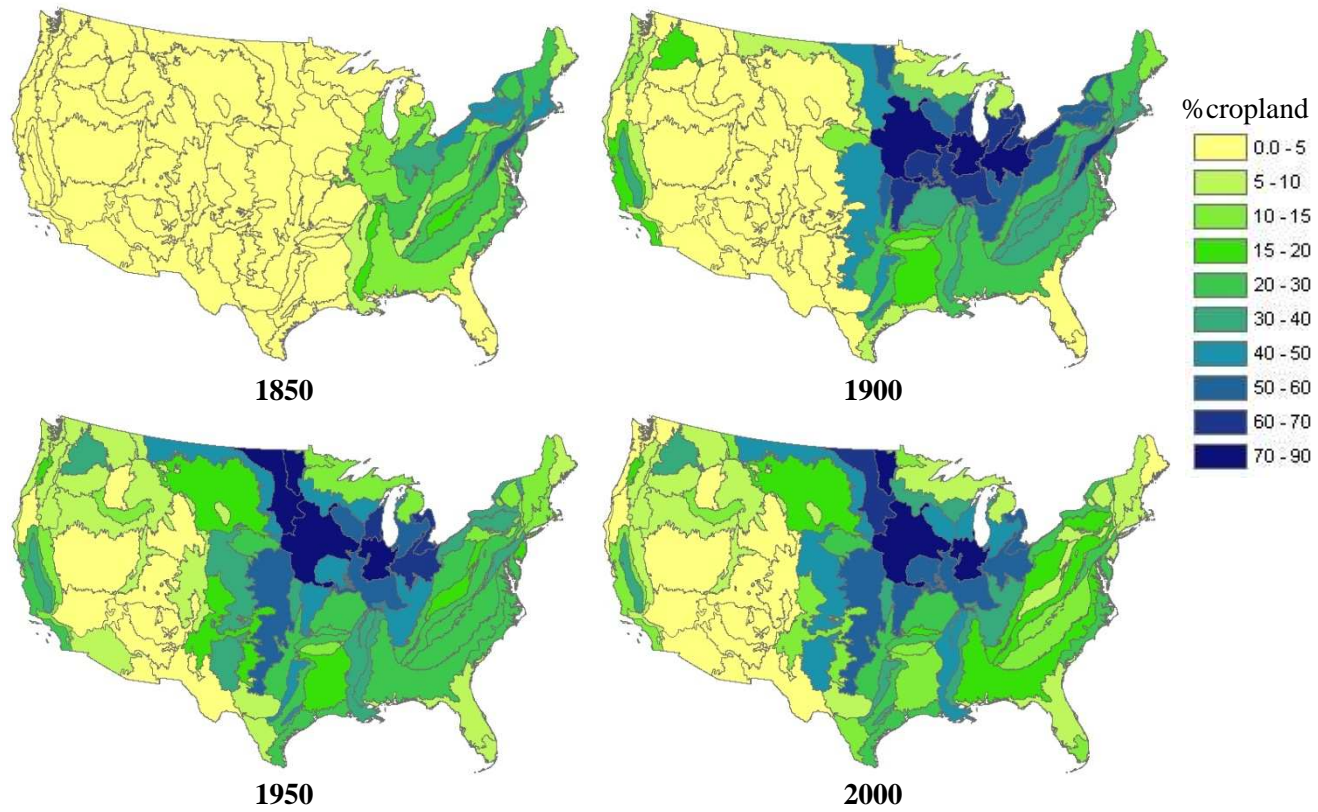
Fig. 4.1: Cropland distribution in 1850, 1900, 1950, and 2000.

4.2 Data and Methods

Data used in the study are listed in Table 4.1. All data are re-gridded to a common resolution of 0.5^O (approximately 50 km) to match with county resolution (approximately 42 km). A GIS (Geographic Information System) program is developed to re-grid county based population density and cropland percentage data (1850 to 2000 per decade). Population and cropland estimates are taken from the United States census data (http://www.nhgis.org; Waisanen and Bliss, 2002). High resolution (1-km) elevation, slope, and topographic index data are re-gridded using the local area averaging function (area_hi2lores) in NCL (NCAR [National Center for Atmospheric Research] Command Language). Dryness index (DI) is an indicator of water potential (water rich versus water poor regions), and is calculated as the ratio of annual evaporative power (net radiation divided by latent heat of vaporization) to precipitation (e.g., DI < 1 means humid region, and DI > 1 means semi-dry to dry region). Annual temperature and precipitation data are based on high resolution (4-km) PRISM climate data (1950 to 2000 average). The biophysical suitability index for cultivation (SUIT) is a function of growing degree days, moisture index, soil carbon content, and pH (Ramankutty et al., 2002). The selection of land-cover change drivers is based on the past literature (see Table 3 in Sohl et al., 2007) and availability of spatial data for the conterminous United States.

An ecoregion denotes a relatively homogeneous area in terms of soil, topography, climate, vegetation and hydrology. Ecoregions are designated as the spatial framework for environmental resource management by the U.S. Environmental Protection Agency (EPA), and this framework has been used in a number of land-cover change studies (e.g. Brown et al. 2005; Loveland et al., 1999; Sohl et al., 2010). There are 84 Level-III ecoregions in the conterminous United States, and ecoregion delineation is obtained from EPA (http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm; Omernik, 1987). Half-degree resolution data is clipped to the ecoregion boundaries and ecoregion average values of all variables are obtained using the local area averaging method. Fig. 4.1 shows ecoregion average %cropland data for 1850, 1900, 1950, and 2000. Ecoregion average values for driver variables are shown in Appendix-A.

Table 4.1: List of dataset used in the study.

Sl. No.	Dataset	Source	Original Resolution	Regridding Method	Final Resolution
1	Population Density (PD)[a]	NHGIS[b]	County (~ 42 km)[h]	area weighted average	0.5 degree (~50 km)
2	Cropland % (CP)[a]	WB2002[c]			
3	Elevation (ELEV)	HYDRO1K[d]	1 km	local area averaging	
4	Slope (SL)				
5	Topographic Index (TI)				
6	Dryness Index (DI)	PRISM[e]	4 km		
7	Annual Temperature (AT)				
8	Crop Suitability Index (SUIT)	RM2002[f]	0.5 degree (~50 km)		
9	Water and Wetland % (WP)	NLCD2001[g]	30 m	Local area sum	
10	Urban land % (UP)				

a- PD and CP estimates are from 1850 to 2000 per decade; b - National Historical Geographic Information System; c - Waisanen and Bliss, 2002; d - USGS topographic datasets derived from 30 arc-second (~ 1 km) digital elevation model; e - Parameter-elevation Regressions on Independent Slopes Model; f - Ramankutty et al., 2002; g - National Land Cover Dataset 2001; h - County resolution (~42 km) is based on average county area (1766 km2) of the conterminous United States.

A multiple linear regression model with stepwise selection method (entry and stay significance level = 0.10) is used to determine major drivers explaining variance in the spatial distribution of %cropland for each decade from 1850-2000. In the stepwise selection method, variables already in the model do not necessarily stay in the model. To get a more meaningful result, only ecoregions having average population density of greater than one person/km^2 are included in the stepwise regression analysis[1]. Power transformation (Box & Cox transformation) is used to bring non-normally distributed variables into a normal distribution. Only one in a pair of independent variables having higher correlation (correlation coefficient > 0.7) is included in stepwise regression, and model residuals are checked for normality using a quantile-quantile plot.

---

[1] This threshold is 10 times threshold used in HYDE3 model for cropland distribution because for HYDE3 model simulation time period is 1700 to 1900, whereas in this study major focus is from 1850 to 2000. We also recognize that this threshold is an arbitrary number.

To describe the cropland trajectory (time series) from 1850 to 2000 a hybrid analytical function (*gamma + beta*) is fitted to the normalized %cropland (Eq. 4.1) for each ecoregion.

$$NCP_{i,y} = \frac{CP_{i,y}}{\max(CP_i)} * 100 \qquad (4.1)$$

$$F(x) = \Gamma(x) + \beta(x) \qquad (4.2)$$

where $NCP_{i,y}$ is the normalized %cropland for year y (1850 to 2000 per decade) in ecoregion i, $CP_{i,y}$ is the %cropland for year y in ecoregion i, and $max(CP_i)$ is the maximum cropland percentage in ecoregion i between year 1850 and 2000 (CP_{max}). First a gamma function is fitted to NCP, then beta function is fitted to absolute value of residual ($NCP(x) – \Gamma(x)$). Gamma and beta function parameters (α and δ for gamma function, and λ and ω for beta function) are optimized by minimizing the sum of squares residuals ($[NCP(x)-\{F(x)\}]^2$). Here, x is decade number (1 to 16) for 1850 to 2000. Goodness of fit is evaluated in terms of Nash-Sutcliffe Efficiency coefficient (NSE). Model formulation details are given in Appendix-B.

Inter-decadal change in %cropland ($CPC_{i,y} = CP_{i,y} – CP_{i, y-10}$) during the 20th century is analyzed against inter-decadal and inter-annual precipitation variability using correlation analysis. Inter-annual precipitation variability is quantified as number of years in a decade having mean standardized departure (absolute values) greater than average mean standardized departure for the 20th century using PRISM data.

4.3 Results

The contributions of different drivers of land-cover change in explaining variance in cropland spatial distribution (partial R^2) are shown in Fig. 4.2. More than 80 percent (average: 87%) variance is explained during the second half of the 19th century, and more than 70% (average: 76%) variance is explained during the 20th century. Population density was the major driver during the 19th century, but its contribution has decreased during the last decades of the 19th century and early half of the 20th century. In the recent decades (1970 to 2000), population density does not explain any variance in cropland spatial distribution. Biophysical suitability for cultivation (SUIT) has shown increasing influence during the 20th century, and has become a

major driver since 1920. The contribution of biophysical suitability has remained major and steady (~60%) in the recent decades. Other minor variables explaining variance in recent decades are elevation (~10%), and annual temperature (4%). Regression coefficient sign (+ / -) shown in Fig. 4.2 are the same for all decades. Further details are given in Appendix-C.

The time series data of %cropland, %urban land, and population density, shown in Fig. 4.2, are average values for the conterminous United States. The %urban land is based on an urban-model for which population density is found an adequate driver (R^2: 0.77; urban-model details are given in Appendix-D). From 1850 to 2000, the population density has increased twelve fold from 2.9 to 34.4 persons/km^2, and % urban land has increased three fold from 1.1% to 3.4%. The %cropland reached its peak (28%) in 1940, and shows a steady decline since then to reach 22% in 2000.
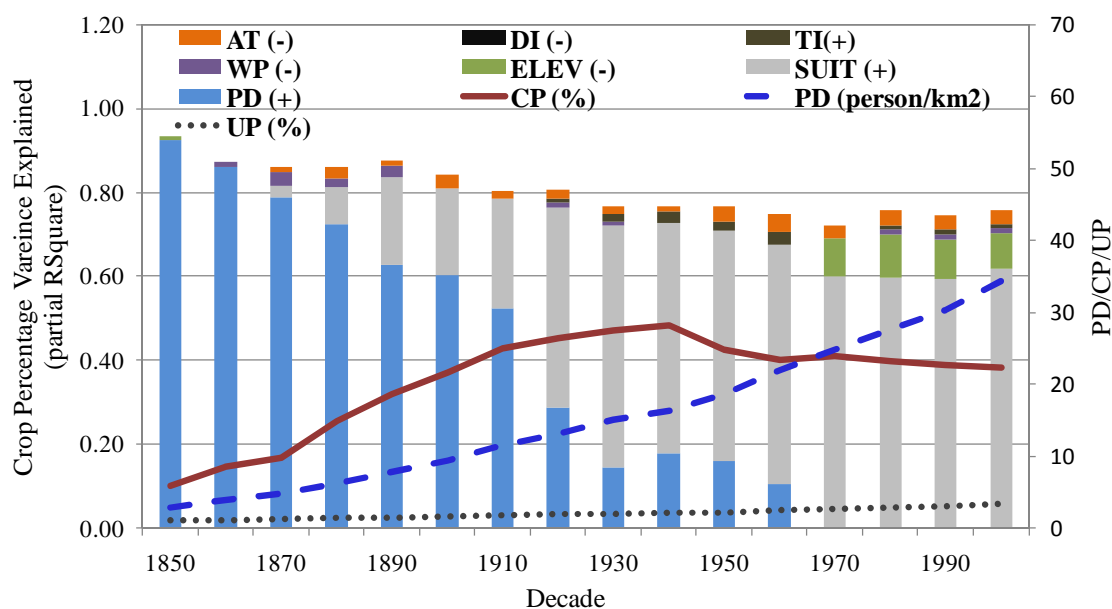
Fig. 4.2: Crop percentage variance explained by socio-economic (PD) and biophysical (SUIT, ELEV, AT, TI, WP, and DI) factors in the conterminous United States. Time series (conterminous USA average) of crop percentage, population density, and urban percentage are also shown.

Fig. 4.3 shows *NCP* time series and fitted hybrid analytical function [*F(x)*] for two example ecoregions (Eco58, and Eco54). These two ecoregions are representative of two major cropland trajectory observed in the data: (1) cropland has continually declined after reaching its peak (e.g. ecoregions in the New England region), referred to as '*declining cropland trajectory*' in the remaining text, and (ii) cropland has stabilized after reaching its peak although at slightly smaller level (e.g. ecoregions in the Corn Belt), referred to as '*stabilized cropland trajectory*' in the remaining text.

The hybrid function adequately describes the *NCP* time series for all ecoregions (average NSE: 0.94). The beta function is the secondary function in the hybrid function, and it describes approximately 19% of total area under the curve. If beta function parameters $\lambda/\omega < 1$ then the beta function compensates for the gamma function residual in the first half of *NCP* time series (Fig. 4.3 (a)). If $\lambda/\omega > 1$ then the beta function compensates for the gamma function residual in the second half of *NCP* time series [Fig. 4.3 (b)]. The scale parameter (δ) of the gamma function is negatively correlated with peak crop percentage year [year (*CP_{max}*), correlation coefficient: -0.54, see Appendix-E]. Higher values of the shape parameter (α) of the gamma function indicate lag start [d(*NCP*)/dt = 0 for initial decades] in the *NCP* time series. A ratio of beta function area to gamma function area from 1940 to 2000 [ArB2G$_{(1940\text{-}2000)}$] indicates compensating effect of beta function over falling limb of gamma function during later half of 20th century. Year (*CP_{max}*) refers to the peak crop percentage year in the data. A table comprising of gamma and beta function parameters, *NSE*, *CP_{max}*, year (*CP_{max}*), ArB2G$_{(1940\text{-}2000)}$, and SUIT for all ecoregions are given in Appendix-E.

(a) Ecoregion 58 [NSE: 0.98, gamma(α: 1.79, δ: 4.07), beta(λ: 1.16, ω: 6.14), beta function area/total area: 0.10, CP_{max}: 36% (1880), SUIT: 0.27]

(b) Ecoregion 54 [NSE = 0.97, gamma(α=2.46, δ=4.25), beta(λ=6.11, ω=1.46), beta function area/total area: 0.18, CP_{max}=83% (1910), SUIT=0.90]
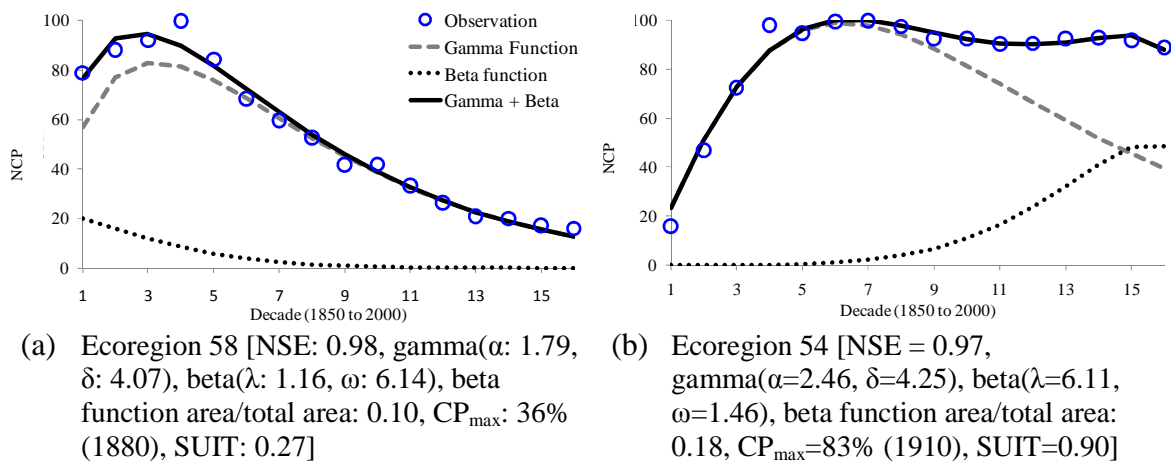
Fig. 4.3: Normalized %cropland time series in ecoregion 58, and 54.

Fig. 4.4(a) shows USA major crop ecoregions (historical) that have $CP_{max} \geq 30\%$ (a total 40 ecoregions). Major crop ecoregions are mostly located in the eastern and central part, and cover 49% of the conterminous United States area. CP_{max} is positively correlated with SUIT index (correlation coefficient: 0.51, also see SP5) in the major Crop region. Out of 40 ecoregions, 28 ecoregions have SUIT ≥ 0.50 (Group A), and the remaining 12 ecoregions have SUIT < 0.50 (Group B). In Group A, 93% ecoregions (26 out of 28) show '*stabilized cropland trajectory*' (average λ/ω: 5.4, average $ArB2G_{(1940\text{-}2000)}$: 0.35); and in Group B, 75% ecoregions (9 out of 12) show '*declining cropland trajectory*' (average λ/ω: 0.31, average $ArB2G_{(1940\text{-}2000)}$: 0.05). Hence, it can be argued that if a region is biophysically suitable for agricultural activity (SUIT > 0.5), then it is more likely that the agricultural activity have been/will be sustained in the region [e.g. Fig. 4.3(b)]. On the other hand, ecoregions where agricultural activity had developed under the influence of population in early decades , even if these regions were not biophysically suitable for agriculture (SUIT < 0.5), have shown continual decline in the cropland % [e.g. Fig. 4.3(a)].

Out of 40 major crop ecoregions (historical), 17 ecoregions are selected for the analysis of agricultural expansion phase (rising limb). In these 17 ecoregions, the hybrid function curve describe rising limb from minimum 30% to maximum 100%. The number of decades taken to go from 30% to its maximum value by the hybrid function curve is termed as the Major Rising Decades (MRD). The minimum cutoff of 30% ensures that: (1) the ecoregion is not developed at the beginning of observation period, and (2) for ecoregions where development started in 20th century a warm up/ lag period (time required to reach 0 to 30%) is taken out of the analysis. For ecoregions that reached their peak in 1950 or earlier (13 out of 17), the expansion phase is found similar (average MRD = 5.2) with a time lag. For ecoregions that reached their peak in the latter half of 20th century slower expansion phase (average MRD = 9.2) is found. A comparison of Eco55 frontier behavior (MRD = 5) with Eco42 (MRD = 6), and Eco73 (MRD = 11) is shown in Fig. 4.5 (also see Table AE.2). Further data are needed to support this finding because there are only 4 major crop ecoregions in the conterminous United States that reached their peak during latter half of the 20th century.
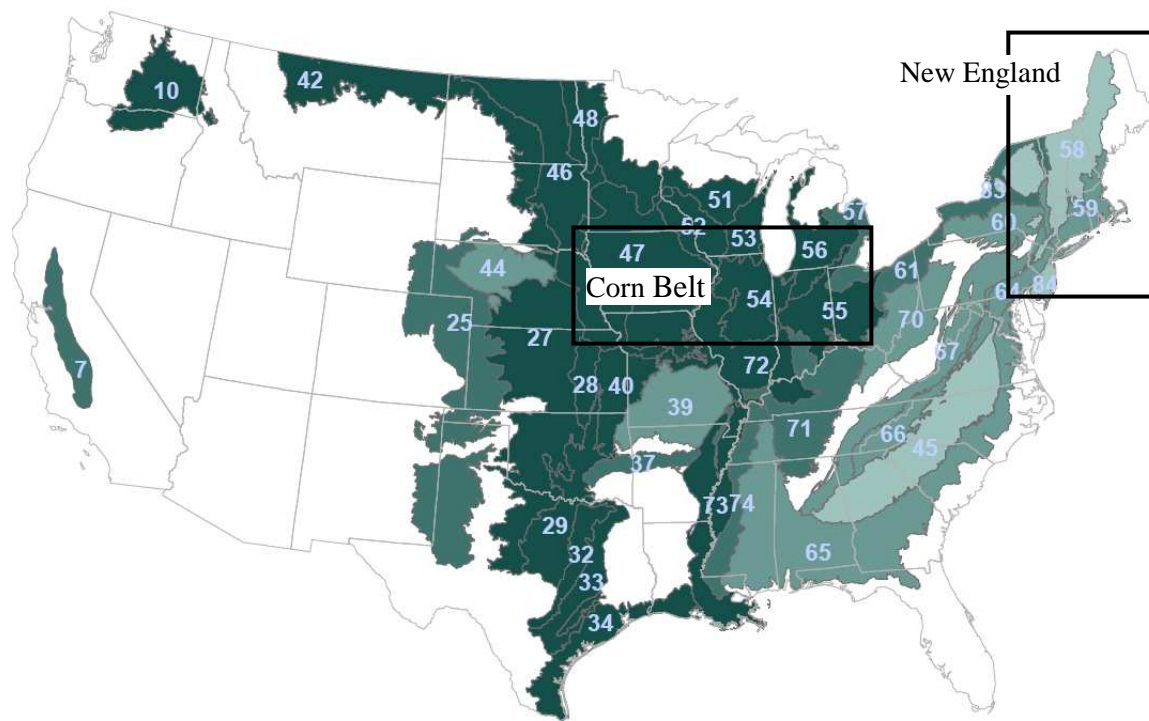
(a)
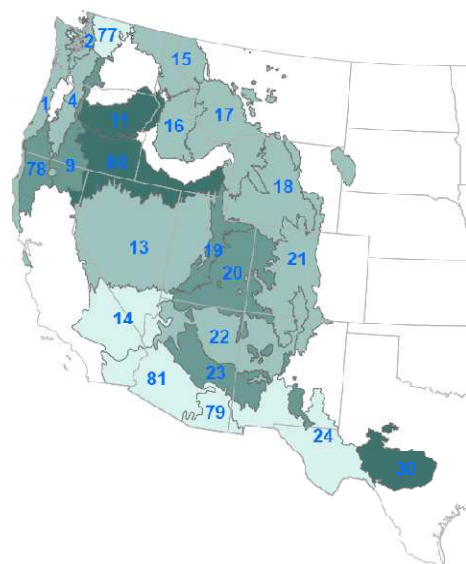
(b)

Fig. 4.4: (a) USA major crop regions (historical, $CP_{max} \geq 30\%$) (b) USA minor crop region ($CP_{max} < 10\%$)
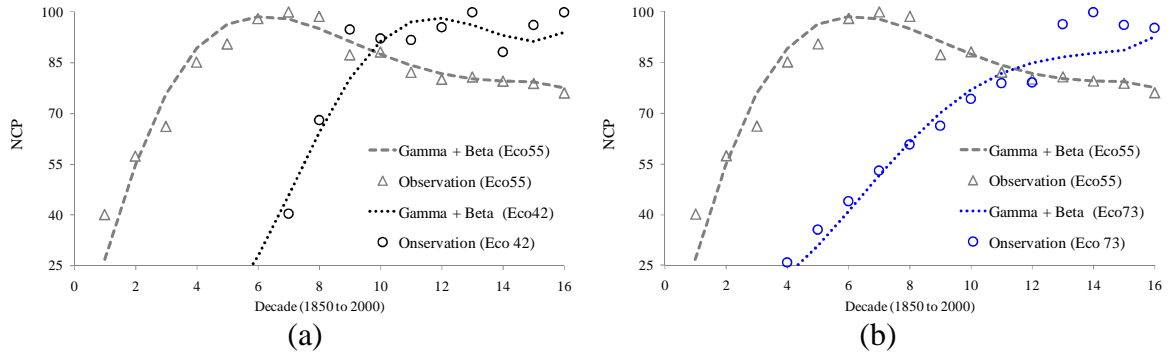
[**Eco55:** (year (CP_{max}): 1910, CP_{max}: 78%, SUIT: 0.92), NSE: 0.90, gamma(α: 2.37, δ: 4.26), beta(λ: 4.62, ω: 1.20); **Eco 42:** year (CP_{max}): 1970, CP_{max}: 45%, SUIT: 0.83; NSE: 0.98, gamma(α: 8.84, δ: 1.40), beta(λ: 5.05, ω: 0.87); **Eco 73**: year (CP_{max}): 1980, CP_{max}: 49%, SUIT: 0.81, NSE: 0.96, gamma(α: 5.23, δ: 2.90), beta(λ: 1.29, ω: 0.71)]

Fig. 4.5: Comparison of Eco55 expansion phase with Eco 42, and Eco73

Twenty three ecoregions have $CP_{max} < 10\%$ (area average CP_{max}: 5%) and all these ecoregions have shown a '*stabilized cropland trajectory*' (average λ/ω: 7.0, and average ArB2G$_{(1940-2000)}$: 0.28), although these regions are not biophysically suitable for agriculture (area average SUIT: 0.33). These ecoregions are located in the western United States and cover 30% of the conterminous United States area (Fig. 4(b)). Hence it can be concluded that very low intensity crop activity ($CP_{max} < 10\%$) has been sustained despite low biophysical suitability of the region. The remaining 20 ecoregions (20.4 % of the conterminous United States area, and area average SUIT index: 0.54) have $20\% < CP_{max} < 10\%$. Only 3 of these 20 ecoregions have shown '*declining cropland trajectory*' (Eco 68, 69, and 82, all located in the eastern United States, see Fig. E.3).

Negative correlation is found between the inter-decadal change in cropland percentage (*CPC*) and decadal average precipitation in major crop regions (Group A ecoregions, 28 ecoregions, average correlation coefficient: -0.25). Seven out of these 28 ecoregions show correlation coefficients less than -0.50 (average:-0.59; Eco34, Eco37, Eco48, Eco51, Eco53, Eco55, and Eco73). Decadal average precipitation and *CPC* for Eco55 are shown in Fig. 4.6. Weaker negative correlation (average: -0.13) is found between *CPC* and inter-annual precipitation variability.
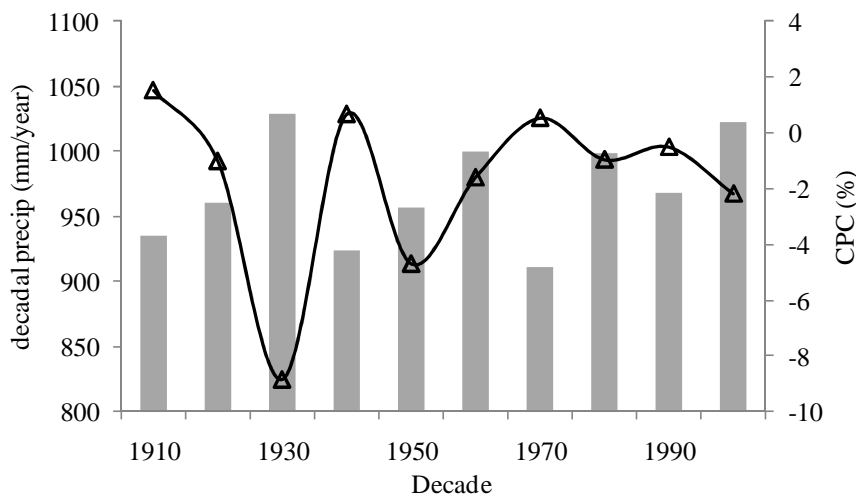
Fig. 4.6: Inter-decadal change in crop percentage (*CPC*) and decadal average precipitation in Eco55.

4.4 Discussion and Conclusion

A quantitative analysis of land-cover change trajectory and its driving forces is presented in this study. The spatial distribution of cropland was governed by population distribution during the 19th century, and biophysical suitability during the 20th century. This finding is consistent with theoretical understanding in the literature; e.g., theory of increasing agricultural adjustment to land quality (Mather and Needle, 1998), and economic development pathway (farming on low suitability land is not economical, and alternative non-farm jobs are present; Rudel et al., 2005). Results of this study have important inferences for USA's near future (2010-2050) landscape including: (1) continued intensification of agricultural activity in biophysically suitable major crop regions (e.g. Corn Belt, Central Plains, Mississippi Valley, and Northern Glaciated Plains); i.e., these regions are not showing sharp decline in cropland as evidenced in the New England region (Foster and Aber, 2004); (2) Despite low biophysical suitability and water constraints, low intensity crop activity ($CP_{max} < 10\%$) is maintained in the western United States. Human intervention in the form of irrigation projects has supported low intensity crop activity in the western United States[2].

---

[2] (a) Of the total water consumptively used for irrigation in the United States, 86% was used in the western United States (west of 100° meridian line) in 1995. (b) Average reservoir capacity in the western United States is approximately 1.5 times the mean annual flow in major western rivers. For Colorado River

The cropland trajectory presented in the study has followed the three general phases of LULC change trajectory unidirectionally. These three phases are: (1) frontier [NCP: 0-30], (2) agricultural expansion [NCP: 30-100], (3) agricultural intensification [NCP: after 100] accompanied by industrialization and urbanization (Mustard et al., 2004). The frontier phase was driven by exogenous factors such as influx of European settlers in New England region and their westward progression (Williams, 1989), and/or political motive to integrate/secure/control new territory (Lambin et al., 2001). The expansion phase is a major natural resource extraction phase by growing population with a perception of unlimited resource availability. However, the expansion phase has been capped by ecoregions biophysical suitability as evident in: (1) positive correlation between CP_{max} and SUIT in major crop region, (2) %cropland in ecoregions of western United States [Fig. 4.4(b)] did not exceed 10%, and cropland trajectory has stabilized.

The intensification phase is driven by both endogenous factors such as biophysical suitability (agricultural adjustment to land quality), as well as exogenous factors such as technological development. What would have been the landscape in the United States without technological development? We do not know. But we do know that the intensification phase has been greatly supported by technological development such as chemical fertilizer and mechanization in agriculture[3], and better transport infrastructure in the post World War II period (Dimitri et al., 2005; Howarth et al., 2002). Between 1948 and 2005, agricultural production has increased 2.66 times without increase in agricultural area (Fuglie et al., 2007). The technological innovations also explain the longer/stretched agricultural expansion (MRD > 5) for the ecoregions in which the peak was reached in the latter half of the 20th century.

The government interventions in the form of price support and supply controls, as well as concerns about adverse ecological impacts from intensive, large-scale farming (e.g., coastal hypoxia, and stream habitat impairment) have also played a minor role in determining cropland trajectory (Dimitri et al., 2005). For example, the total area retired under the Conservation Reserve Programs is 10% of total planted area in 2010 (Cowan, 2010, USDA 2010). However, total area retired in the CRP is approximately 3 times the natural variability (1- standard

reservoir capacity is 5.5 times the mean annual flow. (c) Declining water levels in all major aquifers in the western United States [Anderson, and Woosley, 2005].
[3]Fertilizer input has increased 4 fold from 1961 and 1999 (Howarth et al., 2002), labor input has decreased from 16% of total labor force in 1945 to 1.9% of total labor force in 2000, average farm size has increased from 150 acres (approx.) in 1940 to 450 acres (approx.) in 2000, and mechanization has increased e.g. 2.4 million tractors in 1945 to 4.7 million tractors in 1960, during the same time number of mules and horses used in the farm work has decreased from 11.6 million to 3 million (Dimitri et al., 2005).

deviation) found in the cropland data (1982-2010; Fig. 4.7). In the CRP, which was started in 1985, government pays farmers for keeping the environmentally sensitive land idle or planting cover crops for environmental protection purpose (e.g. erosion control).

This study has presented a foundational work for designing a dynamic (contribution of drivers to change with time) land-cover change (cropland) model for the United States. In some cases, contribution of individual drivers (e.g., dryness index) may not be important, but when they are combined with other driver variables, it can show significant contribution (e.g., SUIT). Future work should include: (1) refinement in SUIT data by incorporating detailed topography and soil data, and effects of irrigation (Ramankutty et al., 2002), (2) consideration of feedback between climate change and biophysical suitability, (3) issue of spatial scale – how to downscale results from ecoregion to county-scale or other appropriate management-scales, and vice-versa, and (4) interaction among eco-regions.
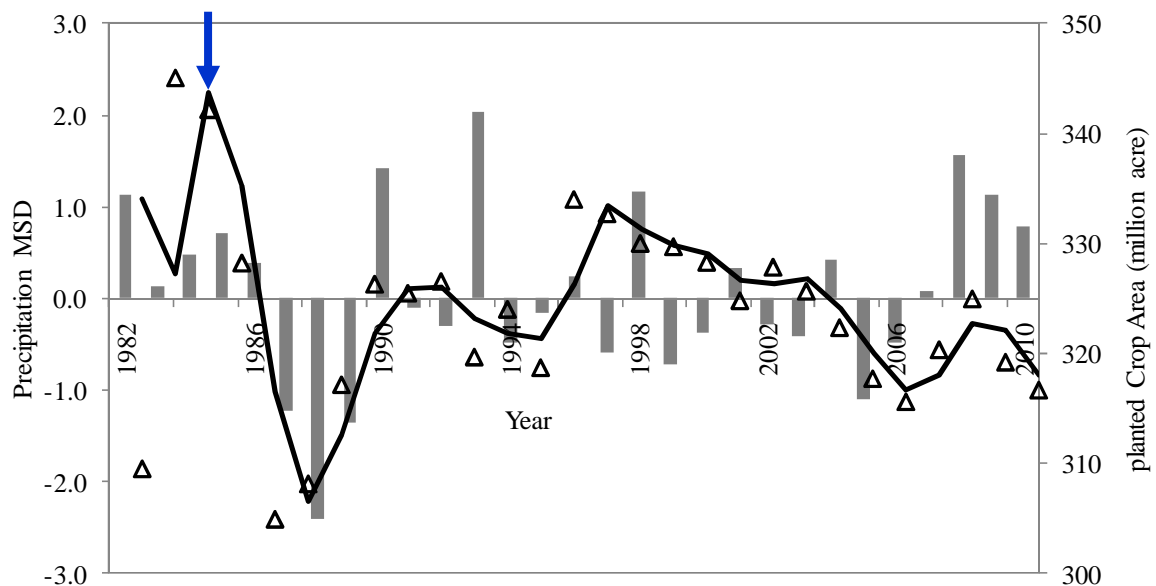
Fig. 4.7: Planted principle crop area in USA from 1982 to 2010 (secondary vertical axis). Interannual climate variability is shown by precipitation mean standardized departure (MSD) for the Midwestern United States (Eco 39-40, 46-49, 51-57, and 71-72). Solid line shows 2 years moving average. Arrow on the curve indicates year 1985 when CRP was implemented.The Midwestern United States has 41% of total cropland in the United States in 2000.

4.5 <u>Summary</u>

Land Use Land Cover (LULC) change is an important anthropogenic-driven phenomenon that has occurred globally over the past several centuries. However, the biophysical and socioeconomic factors contributing towards LULC change over long time periods (decade to century) remain poorly investigated. This study has attempted to relate the historical spatial and temporal (1850-2000) changes in cropland cover in the conterminous United States to several socioeconomic and biophysical drivers using ecoregion based spatial framework. County level cropland and population data, and high resolution topography, climate, and biophysical suitability data are used in the study. The cropland trajectory from 1850 to 2000 for each ecoregion is analyzed using a hybrid analytical function. The spatial distribution of cropland was governed by population distribution during 19th century and biophysical suitability during 20th century. The cropland trajectories show that the United States is in agricultural intensification phase in post world war II period; the agricultural area is increasingly adjusted to the region of high biophysical suitability (for cropland). Low biophysically suitable ecoregions in the Eastern United States have shown continual decline in cropland. The cropland trajectories in high biophysically suitable region, such as Corn Belt, Central Great Plains, lower Mississippi Valley, and Northern Glaciated Plains have been stabilized. Low intensity cropland cover ($< 10\%$) is sustained in the Western United States despite low biophysical suitability of the region. The agricultural intensification phase is expected to continue in the near future (2010-2050) landscape of the United States.

CHAPTER 5: SYNTHESIS

5.1 <u>Hydroclimatology</u>

Hydroclimatology is an emerging science at the interface of hydrology and climatology (atmospheric science). Hydroclimatology emphasizes the role of the climate system in spatial and temporal variations of precipitation minus evapotranspiration difference, and the consequences of this imbalance on a broad range of natural processes including water quantity, quality, and biodiversity (Hirschboeck, 2009; Shelton, 2009, p. 6-8). Roads et al. (2003) has articulated that "*meteorologist and hydrologists had to come together to develop a better understanding of the coupled land atmosphere system at a scale much larger than typically studied by hydrologists and at a scale much smaller than traditionally studied by meteorologists*". The study presented in this dissertation fits into the broader framework of hydroclimatology.

Hydroclimatology utilizes data from the climate system (e.g. temperature, pressure, and humidity) and the terrestrial hydrologic processes (e.g. precipitation, runoff, and soil-moisture) to understand local and regional scale hydroclimatic events (e.g. floods) in the context of large scale atmospheric patterns (Hirschboeck, 1988, 1991; Shelton, 2009, p. 305-343). The hydroclimatology framework is also useful for the examination of human modification to the climate system and the hydrologic cycle (Shelton, 2009, p. 6-8). This study has added some new dimensions to the hydroclimatology framework by: (1) using a coupled climate model to study the long term hydroclimatic changes at regional scale; and (2) seamlessly utilizing historical census data and present day satellite data to study the long term LULC change and its impact on regional hydroclimatology.

Hydrologists have traditionally used outputs from global climate models to study the effects of climate change on water resources (Abbaspour et al., 2009; Milly et al., 2005; Vorosmarty et al., 2000). The importance of LULC change forcing in the regional and global climate system

(Bonan, 1997, 1999; Pielke et al. 2002; Pielke and Niyogi 2010) has moved the boundary from the climate being an external forcing to the climate being an interactive component of the hydrologic landscape. Hence, a coupled land-atmosphere modeling system is needed to assess long term hydroclimatic changes. However, performing coupled climate modeling experiment has rarely been done under the realm of hydrology.

High performance computing infrastructure is needed to perform coupled climate modeling experiments. Although community based supercomputing infrastructure such as TeraGrid has become available in recent years (Basumallik et al., 2007), its use for hydroclimatic studies has remained minimal. In my opinion, two factors contributing to this are: (1) lack of formal graduate training program, such as specially designed course work and/or workshop, and (2) disciplinary resistance, such as going beyond the traditional approach of rainfall-runoff modeling among hydrologists.

5.2 <u>LULC Change versus Climate Change</u>

One important finding presented in Chapter 2 is that the effect of climate change and LULC change can show differently in different seasons. While the climate change has resulted in statistically significant warming in the winter, the wetland drainage has resulted in statistically significant warming in the summer. The regional warming in the summer due to reduced evapotranspiration/latent heat flux is also supported by Lawrence and Chase (2010). This finding is important because IPCC-AR4 has only considered radiative cooling effect of LULC change due to increased surface albedo (Fig. 2.4 in IPCC-AR4 Synthesis Report). Further analysis of the effects of LULC change and the climate change at regional scale can add details to the regional climate change projections.

5.3 <u>Global versus Regional Climate Model</u>

Regional climate models have advantage of higher resolution (~ 50 km) over global climate models (~ 140 km). Regional climate model simulations results are limited by the lateral boundary condition taken from a coarse resolution reanalysis/global climate model outputs (Liang

et al., 2004; Zhu et al., 2009). The approach used in this study that involves the use of a global climate model for regional scale hydroclimatic change study needs further investigation. Specifically, the advantage of dynamical consistency in a global climate modeling experiment (large scale atmospheric circulation sees the LULC change) over a regional climate modeling experiment (given lateral boundary condition which does not see LULC change; e.g. Diffenbaugh, 2009) needs to be ascertained. It is also worth mentioning that global climate model resolution has increased in recent years, for example CCSM4 run is being performed at 50 km resolution.

5.4 Reanalysis Evaluation

The reanalysis evaluation presented in this study was confined to the surface water and energy budget terms. Similar evaluation study incorporating atmospheric water and energy budget terms should be conducted. A basin scale approach and Budyko curve provided useful tools for the reanalysis and climate model evaluations. The AmeriFlux observations were also found useful for model evaluations. As many new reanalysis projects, such as MEERA (Bosilovich, 2008), and CFSR (Saha et al., 2010) have become available in recent years, evaluation tools developed in this study can be useful.

Many studies including this study have found limitations in reanalysis hydrologic outputs (Hagemann et al., 2005; Lucarini et al., 2007). While the major focus in reanalysis projects involve the assimilation of atmospheric observations, the assimilation of surface observation has received less attention. Future reanalysis projects may consider the assimilation of surface observations, more particularly runoff observations/offline model outputs, because runoff will also constrain ET output from a reanalysis model.

The issue of point scale AmeriFlux observations versus climate model grid cell outputs was much debated during the study. This study presented some results related to this issue using long term precipitation and temperature observation from a relatively flat terrain region (Indiana and Illinois). Similar study from a mountainous region can shed more light on this issue. This study also found that the comparison of point scale observation with climate model grid cell outputs can also depend upon the variable of interest, such as latent heat flux versus sensible heat flux. Further investigation on this issue will need a suitably designed field experiment with gridded

network of tower flux sites, and consideration of a moisture advection component among the adjacent grid cells/soil column in the land surface modeling.

5.5 LULC Change Modeling

LULC change modeling work presented in the study (Chapter 4) shows regional and temporal variations in the LULC change trajectory driven by the biophysical and socioeconomic factors. As LULC change modeling continue to evolve (MNP, 2006), historical LULC change trajectory can be a valuable input for future projections. For example, if an ecoregion is showing long term historical declining cropland trajectory, it is less likely that this ecoregion will show an increase in cropland in the near future. Similarly, increasing cropland trajectory has been capped to a particular maximum value, mainly determined by biophysical constraints, and then cropland trajectory either stabilizes or starts declining. The long term persistence behavior in the LULC change trajectory should be explored further.

Lambin and Mayfroidt (2010) suggested two theoretical pathways of land use transition: (1) negative socio-ecological feedback that arises due to endogenous factors such as severely declined natural ecosystem services, and (2) socio-economic dynamics that arise due to exogenous factors such as technological innovation and economic modernization. This study suggests that the cropland trajectory is driven by socio-economic dynamics (stabilized or declining cropland trajectory in post world war II period) within the bound of biophysical constraints (increasing agricultural adjustment to land quality), at least in the latter half of 20th century. This study can be useful to other parts of the world in terms of: (1) maximum cropland for a given biophysical suitability factor, (2) duration and magnitude of the expansion phase, and (3) declining versus stabilized trajectory depending upon the biophysical suitability of an ecoregion.

LIST OF REFERENCES

LIST OF REFERENCES

Abbaspour, K. C., M. Faramarzi, S. S. Ghasemi, and H. Yang (2009), Assessing the impact of climate change on water resources in Iran, Water Resour. Res., 45, W10434, doi:10.1029/2008WR007615

Adegoke, J. O., R. A. Pielke Sr., and A. M. Carleton (2007), Observational and modeling studies of the impacts of agriculture-related land use change on planetary boundary layer processes the central U.S., Agric.For. Meteorol., 142, 203–215, doi:10.1016/j.agrformet.2006.07.013.

Anderson, M T., and Woosley, L. H., Jr., 2005, Water availability for the Western United States-Key scientific challenges: U.S. Geological Survey Circular 1261, 85p.

Armentano, T. V., and E. S. Menges (1986), Patterns of change in the carbon balance of organic soil-wetlands of the temperate zone, J. Ecol., 74, 755–774, doi:10.2307/2260396.

Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem scale carbon dioxide, water vapor, and energy flux densities, Bull. Am. Meteorol. Soc., 82(11), 2415–2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2.

Baldock, D., et al., 2000, The environmental impacts of irrigation in the European Union, Institute for European Environmental Policy, London.

Basu, N. B., et al. (2010b), Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity, Geophys. Res. Lett., 37, L23404, doi:10.1029/2010GL045168.

Basu, N. B., P. S. C. Rao, E. H. Winzeler, S. Kumar, P. Owens, and V. Merwade (2010a), Parsimonious modeling of hydrologic responses in engineered watersheds: Structural heterogeneity versus functional homogeneity, Water Resour. Res., 46, W04501, doi:10.1029/2009WR007803.

Basumallik, A., L. Zhao, C. X. Song, R. L. Sriver, and M. Huber (2007), A Community Climate System Modeling Portal for the TeraGrid, paper presented at TeraGrid 2007 Conference, NSF TeraGrid Project, Madison, Wis.

Berbery, E. H., Y. Luo, K. E. Mitchell, and A. K. Betts (2003), Eta model estimated land surface processes and the hydrologic cycle of the Mississippi basin, J. Geophys. Res., 108(D22), 8852, doi:10.1029/2002JD003192.

Berg, A. A., J. S. Famiglietti, J. P. Walker, and P. R. Houser (2003), Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes, J. Geophys. Res., 108(D16), 4490, doi:10.1029/2002JD003334.

Bogue, M. B. (1951), The swamp land act and wetland utilization in Illinois, 1850–1890, Agric. Hist., 25, 169–180.Bonan, G. B. (1997), Effects of land use on the climate of the United States, Clim. Change, 37, 449–486, doi:10.1023/A:1005305708775.

Bonan G. B. (1997). Effects of land use on the climate of the United States. Climate Change, 37, 449-486.

Bonan, G. B. (1999), Frost followed the plow: Impacts of deforestation on the climate of the United States, Ecol. Appl., 9(4), 1305–1315, doi:10.1890/1051-0761(1999)009[1305:FFTPIO]2.0.CO;2.

Bonan, G. B. (2001), Observational evidence for reduction of daily maximum temperature by croplands in the Midwest United States, J. Clim., 14, 2430–2442, doi:10.1175/1520-0442(2001)014<2430:OEFROD>2.0.CO;2.

Bonan, G. B. (2008), Ecological Climatology Concepts and Applications, 2nd ed., Cambridge Univ. Press, New York.

Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson, and Z. L. Yang (2002), The land surface climatology of the community land model coupled to the NCAR community climate model, J. Clim., 15, 3123–3149, doi:10.1175/1520-0442(2002)015<3123:TLSCOT>2.0.CO;2.

Bosilovich, M. G. (2008), NASA's Modern Era Retrospective-analysis for Research and Application: Integrating Earth Observations, <http://www.earthzine.org>

Boville, B. A., P. J. Rasch, J. J. Hack, and J. R. McCaa (2006), Representation of clouds and precipitation processes in the Community Atmospheric Model (CAM3), J. Clim., 19, 2184–2198, doi:10.1175/JCLI3749.1.

Breiman, L., J. H. Friedman, R. A. Olshend, and C. J. Stone (1984), Classification and Regression Trees, 358 pp., Wadsworth Int. Group, Belmont, Calif.

Brovkin V., S. Sitch, W. V. Bloh, M. Claussen, E. Bauer, and W. Cramer, 2004. Role of land cover change for atmospheric CO_2 increase and climate change during the last 150 years. Global Change Biology, 10, 1253 – 1266, doi: 10.1111/j.1365-2486.2004.00812.x

Brown, D. G., K. M. Johnson, T. R. Loveland, and D. M. Theobald (2005), Rural land-use trends in the conterminous United States, 1950-2000, Ecological Applications, 15, 6, 1852-1863.

Budyko, M. I. (1958), The Heat Balance of the Earth's Surface, US Dep. of Commer., Washington, D. C.

Bureau of the Census (1932), Fifteenth census of the United States: 1930, Drainage of agricultural lands, report, Washington, D. C. Bureau of the Census (1952), U.S. census of agriculture: 1950, vol. IV,

Bureau of the Census (1973), 1969 census of agriculture, vol. VI, Drainage of agricultural lands, report, Washington, D. C.

Bureau of the Census (1981), 1978 census of agriculture, special report, vol. 5, Drainage of agricultural lands, report, Washington, D. C.

Carlisle, D. M., D. M. Wolock, and M. R Meador (2010). Alteration of streamflow magnitudes and potential ecological consequences: a multiregional assessment. Frontiers in Ecology and the Environment (e-View) doi:10.1890/100053

Carrington, D. P., R. G. Gallimore, and J. E. Kutzbach (2001), Climate sensitivity to wetlands and wetland vegetation in mid-Holocene North Africa, Clim. Dyn., 17, 151–157, doi:10.1007/s003820000099.

Chen, F., and J. Dudhia (2001), Coupling an advanced land surface hydrology model with the Penn State-NCAR MM5 modeling system. Part 1: Model implementation and sensitivity, Mon. Weather Rev., 129, 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Chow, V. T., D. R. Maidment, and L. W. Mays (1988), Applied Hydrology, McGraw Hill, New York.

CIMP5, Lead Authors, K. L. Taylor, R. J. Stouffer, and G. A. Meehl, 2009. A summary of the CMIP5 experiment design. http://cmip-pcmdi.llnl.gov/cmip5/index.html?submenuheader=0

Collins, W. D., et al. (2006), The Community Climate System Model Version 3(CCSM3), J. Clim., 19, 2122–2143, doi:10.1175/JCLI3761.1.

Covey, C., K. M. Achuta Rao, U. Cubasch, P. Jones, S. J. Lambert, M. E. Mann, T. J. Phillips, and K. E. Taylor (2003), An overview of results from the Coupled Model Intercomparison Project, Global Planet. Change, 37, 103–133, doi:10.1016/S0921-8181(02)00193-5.

Cowan, T., 2010, Conservation Reserve Program: Status and Current Issues, Congressional Research Service, 7-5700, www.crs.gov, RS21613

Dahl, T. E. (1990), Wetlands Loss in the United States 1780's to 1980's, Fish and Wildlife Serv., U.S. Dep. of the Inter., Washington, D. C.

Daly, C., G. H. Taylor, W. P. Gibson, T. Parzybok, G. L. Johnson, and P. A. Pasteris (1998), Development of high-quality spatial datasets for the United States, paper presented at First International Conference on Geospatial Information in Agriculture and Forestry, ERIM, Lake Buena Vista, Fla.

Daly, C., G. Taylor, and W. Gibson (1997), The prism approach to mapping precipitation and temperature, paper presented at 10th Conference on Applied Climatology, Am. Meteorol. Soc., Reno, Nev.

Dickinson, R. E., K. W. Oleson, G. Bonan, F. Hoffman, P. Thornton, M. Vertenstein, Z. L. Yang, and X. Zeng (2006), The community land model and its climate statistics as a component of community climate system model, J. Clim., 19, 2302–2324, doi:10.1175/JCLI3742.1.

Dideriksen, R. I., A. R. Hidlebaugh, and K. O. Schmude (1978), Wet soils for crop production in the United States, in Wetland Functions and Values: The State of our Understanding. Proceedings of the National Symposium on Wetlands, edited by P. E. Greeson, J. R. Clark, and J. E. Clark, pp. 632–641, Am. Water Resour. Assoc., Middleburg, Va.

Diffenbaugh, N. S. (2009), Influence of modern land cover on the climate of the United States, Clim. Dyn., 33, 945–958, doi:10.1007/s00382-009-0566-z.

Dimitri, C., A. Effland, and N. Conklin, 2005, The 20th century transformation of U. S. Agriculture and Farm Policy, Economic Information Bulletin Number 3, USDA-Economic Research Service.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006), GSWP-2 multimodel analysis and implications for our perception of the land surface, Bull. Am. Meteorol. Soc., 87(10), 1381–1397, doi:10.1175/BAMS-87-10-1381.

Dominguez, F., and P. Kumar (2008), Precipitation recycling variability and eco-climatological stability—A study using NARR data. Part I:Central U.S. plains ecoregion, J. Clim., 21, 5165–5186, doi:10.1175/2008JCLI1756.1.

Dominguez, F., P. Kumar, and E. R. Vivoni (2008), Precipitation recycling variability and eco-climatological stability—A study using NARR data. Part II: North American monsoon region, J. Clim., 21, 5187–5203,doi:10.1175/2008JCLI1760.1.

Drainage of agricultural land, report, Washington, D. C.Bureau of the Census (1961), U.S. census of agriculture: 1959. final report, vol. IV, Drainage of agricultural lands, report, Washington, D. C.

Eastman, J. L., M. B. Coughenour, and R. A. Pielke (2001), The effects of CO2 and landscape change using a coupled plant and meteorological model, Global Change Biol., 7, 797–815, doi:10.1046/j.1354-1013.2001.00411.x.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren,G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108 (D22), 8851, doi:10.1029/2002JD003296.

Eltahir, E. A. B. (1998), A soil moisture–rainfall feedback mechanism: 1. Theory and observations, Water Resour. Res., 34(4), 765–776,doi:10.1029/97WR03499.

Fall, S., D. Niyogi, R. A. Pielke Sr., A. Gluhovsky, E. Kalnay, and G. Rochon (2010), Impacts of land use land cover on temperature trends over the continental United States: Assessment using the North American Regional Reanalysis, Int. J. Climatol., 30, 1980–1993, doi:10.1002/joc.1996.

Fan, Y., G. Miguez-Macho, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 1.Water table observations and equilibrium water table simulations, J. Geophys. Res., 112, D10125, doi:10.1029/2006JD008111.

Feddema, J. J., K. W. Oleson, G. B. Bonan, L. O. Mearns, L. E. Buja, G. A. Meehl, and W. M. Washington (2005), The importance of land-cover change in simulating future climates, Science, 310, 1674–1678.

Fekete, B. M., C. J. Vörösmarty, and W. Grabs (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, Global Biogeochem. Cycles, 16(3), 1042, doi:10.1029/1999GB001254.

Fenneman, N. M., and D. W. Johnson (1946), Physiographic Division of the Conterminous U.S., U.S. Geol. Surv., Washington, D. C.

Findell, K. L., A. J. Pitman, M. H. England, and P. J. Pegion (2009), Regional and global impacts of land cover and sea surface temperature anomalies, J. Clim., 22, 3248–3269, doi:10.1175/2008JCLI2580.1.

Findell, K. L., and E. A. B. Eltahir (1997), An analysis of the soil moisture–rainfall feedback, based on direct observations from Illinois, Water Resour. Res., 33(4), 725–735, doi:10.1029/96WR03756.

Findell, K. L., and E. A. B. Eltahir (2003), Atmospheric controls on soil moisture-boundary layer interactions. Part I: Framework development, J. Hydrometeorol., 4, 552–568, doi:10.1175/1525-7541(2003) 004<0552:ACOSML>2.0.CO;2.

Foken, T. (2008), The energy balance closure problem: An overview, Ecol. Appl., 18(6), 1351–1367, doi:10.1890/06-0922.1.

Foster, D. R., and Aber, J. D. (Editors), 2004, Forest in Time: The Environmental Consequences of 1,000 years of change in New England, Yale University Press.

Fry, J. A., M. J. Coan, C. G. Homer, D. K. Meyer, and J. D. Wickham (2009), Completion of the National Land Cover Database (NLCD) 1992–2001 Land Cover Change Retrofit product, U.S. Geol. Surv. Open File Rep., 2008–1379, 18 pp.

Fuglie, K. O., J. M. MacDonald, and E. Ball (2007), Productivity growth in U.S. agriculture, Economic Brief Number 9, USDA-Economic Research Service.

Gates, W. L., et al. (1999), An overview of the results of the atmospheric model intercomparsion project (AMIP I), Bull. Am. Meteorol. Soc., 80(1), 29–55, doi:10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2.

Gent, P. R. (2006), Preface to special issue on Community Climate System Model (CCSM), J. Clim., 19, 2121, doi:10.1175/JCLI9020.1.

Giorgi, F., L. O. Mearns, C. Shields, and L. Mayer (1996), A regional model study of the importance of local versus remote controls of the 1988 drought and the 1993 flood over the central United States, J. Clim., 9, 1150–1162, doi:10.1175/1520-0442(1996)009<1150:ARMSOT>2.0. CO;2.

Goldewijk, K. K. and G.Van Drecht (2006), HYDE 3: Current and historical population and land cover, in: MNP (2006) (Edited by A.F. Bouwman, T. Kram and K. Klein Goldewijk), Integrated modelling of global environmental change. An overview of IMAGE 2.4. Netherlands Environmental Assessment Agency (MNP), Bilthoven, The Netherlands, 93-112.

Goolsby, D. A., and W. A. Battaglin (2001), Long-term changes in concentrations and flux of nitrogen in the Mississippi river basin, USA, Hydrol. Processes, 15, 1209–1226, doi:10.1002/hyp.210.

Gordon L. J., W. Steffen, B. F. Jonsson, C. Folke, M. Falkenmark, and A. Johannessen, 2005. Human modification of global water vapor flows from the land surface. Proceeding of National Academy of Science, 102, 21, 7612-7617.

Gordon, L., M. Dunlop, and B. Foran (2003), Land cover change and water vapour flows: Learning from Australia, Philos. Trans. R. Soc., Ser. B, 358, 1973–1984, doi:10.1098/rstb.2003.1381.

Gosain, A. K., A. Rao, and D. Basuray (2006). Climate change impact assessment on hydrology of Indian River basins. Current Science, 90(3), 346-353.

Gruber, N., and J. N. Galloway (2008), An earth-system perspective of the global nitrogen cycle, Nature, 45, doi:10.1038/nature06592.

Hagemann, S., K. Arpe, and L. Bengtsson (2005), Validation of the hydrological cycle of ERA-40, ERA-40 Proj. Rep. Ser. 24, Eur. Cent. for Medium-Range Weather Forecasts, Reading, U. K.

Han, J., and J. O. Roads (2004), U.S. climate sensitivity simulated with the NCEP regional spectral model, Clim. Change, 62, 115–154, doi:10.1023/B:CLIM.0000013675.66917.15.

Herget, J. E. (1978), Taming the environment: The drainage district in Illinois, J. Ill. State Hist. Soc., 71, 107–118.

Hirschboeck, K. K. (1988), Flood hydroclimatology. In Flood Geomorphology, ed. V. R. Baker, R.C. Kochel and P.C. Patton, New York, NY:John Wiley and Sons, pp. 27-49.

Hirschboeck, K. K. (1991), Climate and Floods. In National Water Summary 1988-89 Hydrologic events and Floods and Droughts. USGS water-supply paper 2375, Wahington, D.C.: U.S. government printing office, pp. 67-88.

Hirschboeck, K. K. (2009), Future hydroclimatology and the research challenges of a post-stationary world, UCWR, Journal of Contemporary Water Research and Education, 142, 4-9.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan (2004), Development of a 2001 National Land-Cover Database for the United States, Photogramm. Eng. Remote Sens., 70(7), 829–840.

Howarth, R. W., E. W. Boyer, W. J. Pabich, and J. N. Galloway, 2002, Nitrogen use in the united states from 1961-2000 and potential future trends, AMBIO: A Journal of the Human Environment, 31, 2, 88-96.

Hudson, J. (2007), Physical geography, in The American Midwest, An Interpretive Encyclopedia, edited by R. Sisson, C. Zacher, and A. Cayton, pp. 129–133, Indiana Univ. Press, Indianapolis.

Hurkmans, R. T. W. L., W. Terink, R. Uijlenhoet, E. J. Moores, P. A. Torch, and P. H. Verburg, 2009. Effects of land use changes on streamflow generation in the Rhine basin. Water Resources Research, 45, W06405, doi:10.1029/2008WR007574.

IPCC-AR4 Synthesis Report, Bernstein, L., et al., (2008), Climate Change 2007: Synthesis Report, An Assessment of the Intergovernmental Panel on Climate Change.

Jenkins, D. G., S. Grissom, and K. Miller (2003), Consequences of prairie wetland drainage for crustacean biodiversity and metapopulations, Conserv. Biol., 17(1), 158–167, doi:10.1046/j.1523-1739.2003.01450.x.

Jones, J. R., B. P. Borofka, and R. W. Bachmann (1976), Factors affecting nutrient loads in some Iowa streams, Water Res., 10, 117–122.

Kalnay, E., et al. (1996), The NCEP/NCAR 40‑Year Reanalysis Project, Bull. Am. Meteorol. Soc., 77(3), 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kanamitsu, M., et al. (2002), NCEP/DE AMIP‑II Reanalysis (R-2), Bull. Am. Meteorol. Soc., 83(11), 1631–1643, doi:10.1175/BAMS-83-11-1631(2002)083<1631:NAR>2.3.CO;2.

Karl, T. R., C. N. Williams Jr., P. J. Young, and W. M. Wendland (1986), A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States, J. Clim. Appl. Meteorol., 25, 145–160, doi:10.1175/1520-0450(1986) 025<0145:AMTETT>2.0.CO;2.

Kiehl, J. T., and P. R. Gent (2004), The community climate system model, version 2, J. Clim., 17, 3666–3682, doi:10.1175/1520-0442(2004) 017<3666:TCCSMV>2.0.CO;2.

Kiehl, J. T., C. A. Shields, J. J. Hack, and W. D. Collins (2006), The climate sensitivity of the community climate system model version 3(CCSM3), J. Clim., 19, 2584–2596, doi:10.1175/JCLI3747.1.

Kitanidis, P. K. (1997), Introduction to Geostatistics: Application in Hydrogeology, Cambridge Univ. Press, Cambridge, U. K., doi:10.1017/CBO9780511626166.

Koster, R. D., and M. J. Suarez (1999), A simple framework for examining the interannual variability of land surface moisture fluxes, J. Clim., 12, 1911–1917, doi:10.1175/1520-0442(1999)012<1911:ASFFET>2.0.CO;2.

Koster, R. D., et al. (2004), Regions of strong coupling between soil moisture and precipitation, Science, 305, 1138–1140, doi:10.1126/science.1100217.

Koster, R. D., M. J. Suarez, and M. Heiser (2000), Variance and predictability of precipitation at seasonal‑to‑interannual timescales, J. Hydrometeorol., 1, 26–46, doi:10.1175/1525-7541(2000)001<0026:VAPOPA>2.0.CO;2.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006), World map of the Köppen‑Geiger climate classification updated, Meteorol. Z., 15(3), 259–263, doi:10.1127/0941-2948/2006/0130.

Krysanova V., F. Hattermann, and A. Habeck, 2005. Expected changes in water resources availability and water quality with respect to climate change in the Elbe River basin (Germany). Nordic Hydrology, 36, 321-333.

Kumar, S., and V. Merwade (2009), Evaluation of evapotranspiration component in global climate model, Eos Trans. AGU, 90(52), Fall Meet.Suppl., Abstract H51B-0767.

Kumar, S., and V. Merwade (2011), Evaluation of NARR and CLM3.5 outputs for surface water and energy budgets in the Mississippi River Basin, J. Geophys. Res., 116, D08115, doi:10.1029/2010JD014909.

Kumar, S., V. Merwade, J. Kam, and K. Thurner (2009), Streamflow trends in Indiana: Effects of long term persistence, precipitation and subsurface drains, J. Hydrol., 374, 171–183, doi:10.1016/j.hydrol.2009.06.012.

Kumar, S., V. Merwade, W. Lee, L. Zhao, and C. Song (2010), Hydroclimatological impact of century long drainage in midwestern United States: CCSM sensitivity experiments, J. Geophys. Res., 115, D14105, doi:10.1029/2009JD013228.

Lambert, S. J., and G. J. Boer (2001), CMIP1 evaluation and intercomparison of coupled climate models, Clim. Dyn., 17, 83–106, doi:10.1007/PL00013736.

Lambin and Meyfroidt (2010), Land use transition: Socio-ecological feedback versus socio-economic change, Land Use Policy, 27, 108-118

Law, R. M., et al. (2009), AmeriFlux Level 4 data, http://cdiac.ornl.gov/ftp/ameriflux/data, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab.,Oak Ridge, Tenn.

Lawrence P. J., and T. N. Chase (2010): Investigating the climate impacts of global land cover change in the community climate system model. International Journal of Climatology, DOI: 10.1002/joc.2061.

Lawrence, D. M., P. E. Thornton, K. W. Oleson, and G. B. Bonan (2007), The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmospheric interaction, J. Hydrometeorol., 8, 862–880, doi:10.1175/JHM596.1.

Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res., 112, G01023, doi:10.1029/2006JG000168.

Li, S.-G., C.-T. Lai, G. Lee, S. Shimoda, T. Yokoyama, A. Higuchi, and T. Oikawa (2005), Evapotranspiration from a wet temperate grassland and its sensitivity to micro environmental variables, Hydrol. Processes, 19, 517–532, doi:10.1002/hyp.5673.

Li, Z., and N. Molders (2008), Interaction of impacts of doubling CO_2 and changing regional land-cover on evaporation, precipitation, and runoff at global and regional scales, Int. J. Climatol., 28, 1653–1679, doi:10.1002/joc.1666.

Li, Z., U. S. Bhatt, and N. Molders (2008), Impact of doubled CO_2 on the interaction between the global and regional water cycles in four study regions, Clim. Dyn., 30, 255–275, doi:10.1007/s00382-007-0283-4.

Liang, X.-Z., J. Pan, J. Zhu, K. E. Kunkel, J. X. L. Wang, and A. Dai (2006), Regional climate model downscaling of the U.S. summer climate and future change, J. Geophys. Res., 111, D10108, doi:10.1029/2005JD006685.

Liang, X.-Z., L. Li, K. E. Kunkel, M. Ting, and J. X. L. Wang (2004), Regional climate model simulation of U.S. precipitation during 1982–2002. Part 1: Annual cycle, J. Clim., 17, 3510–3528, doi:10.1175/1520-0442(2004)017<3510:RCMSOU>2.0.CO;2.

Loveland et al., 1999, Land Cover Trends: Rates, Causes, and Consequences of Late-Twentieth Century U.S. Land Cover Change, Research Plan, EPA-IAG Project No. DW14938108-01-0

Lucarini, V., R. Danihlik, I. Kriegerova, and A. Speranza (2007), Does the Danube exist? Versions of reality given by various regional climate models and climatological data sets, J. Geophys. Res., 112, D13103, doi:10.1029/2006JD008360.

Marshall, C. H., Jr., R. A. Pielke Sr., L. T. Steyaert, and D. A. Willard (2004), The impact of anthropogenic land-cover change on the Florida peninsula sea breezes and warm season sensible weather, Mon. Weather Rev., 132, 28–52, doi:10.1175/1520-0493(2004)132<0028:TIOALC>2.0.CO;2.

Mather, A. S., and Needle, C. L., 1998, The Forest Transition: A theoretical Basis, Area, 30, 2, 117-124.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, J. Clim., 15, 3237–3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.

Maurer, E. P., G. M. O'Donnell, D. P. Lettenmaier, and J. O. Roads (2001), Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalysis using an off-line hydrologic model, J. Geophys. Res., 106(D16), 17,841–17,862, doi:10.1029/2000JD900828.

Mayer, A. H. (1935), The Kankakee marsh of northern Indiana and Illinois, Pap. Mich. Acad. Sci. Arts Lett., 21, 359–396. McCorvie, M., and C. L. Lant (1993), Drainage district formation and the loss of midwestern wetlands, 1850–1930, Agric. Hist., 67(4), 13–39.

Menne, M. J., C. N. Williams Jr., and R. S. Vose (2009), United States Historical Climatology Network (USHCN) Version 2 Serial Monthly Dataset, http://cdiac.ornl.gov/epubs/ndp/ushcn/monthly_doc.html, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Mesinger, F. (2000), Numerical methods: The Arakawa approach horizontal grid, global, and limited-area modeling, in General Circulation Model Development, Int. Geophys. Ser., vol. 70, edited by D. A. Randall, pp. 373–419, Academic, San Diego, Calif.

Mesinger, F., et al. (2006), North American Regional Reanalysis, Bull. Am. Meteorol. Soc., 87, 343–360, doi:10.1175/BAMS-87-3-343.

Miguez-Macho, G., Y. Fan, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation, J. Geophys. Res., 112, D13108, doi:10.1029/2006JD008112.

Miller, D. A., and R. A. White (1998), A conterminous United States multilayer soil characteristics data set for regional climate and hydrology modeling, Earth Interact., 2(2), 1–26.

Miller, I., and M. Miller (2004), John E. Freund's Mathematical Statistics With Applications, 7th ed., Prentice-Hall, Englewood Cliffs, N. J.

Milly P. C. D., K. A. Dunne, and A. V. Vecchia (2005). Global pattern of trends in streamflow and water availability in a changing climate. Nature, 438, 347-350.

Milly, P. C. D. (1994), Climate, soil water storage, and the average annual water balance, Water Resour. Res., 30(7), 2143–2156, doi:10.1029/94WR00586.

Mimikou M. A., E. Baltas, E. Varanou, and K. Pantazis (2000). Regional impact of climate change on water resources quantity and quality indicators. Journal of Hydrology, 234, 95-109.

MNP (2006) (Edited by A.F. Bouwman, T. Kram and K. Klein Goldewijk), Integrated modelling of global environmental change. An overview of IMAGE 2.4. Netherlands Environmental Assessment Agency (MNP), Bilthoven, The Netherlands.

Monson, R. K., J. P. Sparks, T. N. Rosenstiel, L. E. Scott-Denton, T. E. Huxman, P. C. Harley, A. A. Turnipseed, S. P. Burns, B. Backlund, and J. Hu (2005), Climatic influences on net ecosystem CO_2 exchange during the transition from wintertime carbon source to springtime carbon sink in a high-elevation, subalpine forest, Oecologia, 146(1), 130–147, doi:10.1007/s00442-005-0169-2.

Monteith, J. L. (1965), Evaporation and environment, in The State and Movement of Water in Living Organisms, edited by G. E. Fogg, pp. 205–234, Cambridge Univ. Press, Cambridge, U. K.

Mustard, J. F., R. S. DeFries, T. R. Fisher, and E. F. Moran, 2004, Land-Use and Land-Cover change pathways and Impacts, Land Change Science, 6, 411-429, DOI: 10.1007/978-1-4020-2562-4_24

Niu, G.-Y., and Z.-L. Yang (2006), Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale, J. Hydrometeorol., 7, 937–952, doi:10.1175/JHM538.1.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res., 110, D21106, doi:10.1029/2005JD006111.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007),Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res., 112, D07103, doi:10.1029/2006JD007522.

Novitzki, R. P. (1978), Hydrologic Characteristics of Wisconsin's wetlands and their influence on floods, streamflow, and sediment, in Wetland Functions and Values: The State of Our Understanding. Proceeding of the National Symposium on Wetlands, edited by P. E. Greeson, J. R. Clark, and J. E. Clark, pp. 377–388, Am. Water Resour. Assoc., Middleburg, Va.

Oleson, K. W., et al. (2004), Technical description of the community land model (CLM), NCAR Tech. Note NCAR/TN-461+STR, Natl. Cent. For Atmos. Res., Boulder, Colo.

Oleson, K. W., et al. (2008), Improvements to the Community Land Model and their impact on the hydrological cycle, J. Geophys. Res., 113, G01021, doi:10.1029/2007JG000563.

Omernik, J.M. (1987), Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the Association of American Geographers 77(1):118-125.

Pastore, C. L., et al. (2010), Tapping Environmental History to recreate America's colonial hydrology, Environ. Sci. Technol., 44, 8798-8803

Paull, R. K., and R. A. Paull (1977), Geology of Wisconsin and Upper Michigan Including Parts of Adjacent States, Kendall-Hunt, Dubuque, Iowa.

Pavelis, G. A. (Ed.) (1987), Farm drainage in the United States: History status and prospects, Misc. Publ. U.S. Dep. Agric., 1455, 126–153.

Phillips, N. A. (1957), A coordinate system having some special advantage for numerical forecasting, J. Meteorol., 14, 184–185.

Pielke Sr. R.A., and D. Niyogi (2010): The role of landscape processes within the climate system. In: J.-C. Otto, R. Dikau (eds.), Landform . Structure, Evolution, Process Control, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_5, Springer-Verlag Berlin Heidelberg, 67-8

Pielke Sr., R.A., G. Marland, R.A. Betts, T.N. Chase, J.L. Eastman, J.O. Niles, D. Niyogi, and S. Running (2002): The influence of land-use change and landscape dynamics on the climate system- relevance to climate change policy beyond the radiative effect of greenhouse gases. Phil. Trans. A. Special Theme Issue, 360, 1705-1719

Pielke, R. A., Sr. (2005), Land use and climate change, Science, 310, 1625–1626, doi:10.1126/science.1120529.

Pongratz, J., C. Reick, T. Raddatz, and M. Claussen (2008), A reconstruction of global agricultural areas and land cover for the last millennium, Global Biogeochem. Cycles, 22, GB3018, doi:10.1029/2007GB003153.

Postel, S. L., G. C. Daily, and P. R. Ehrlich (1996), Human appropriation of renewable fresh water, Science, 271(5250), 785–788, doi:10.1126/science.271.5250.785.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, J. Hydrometeorol., 7, 953–975, doi:10.1175/JHM540.1.

Ramankutty N., and J. A. Foley (1999). Estimating historical changes n global land cover: croplands from 1700 to 1992. Global Biochemical Cycles, 13, 4, 997-1027.

Ramankutty N., J. A. Foley, J. Norman, and K. Mcsweeney, 2002, The global distribution of cultivable lands: current patterns and sensitivity to possible climate change, Global Ecology & Biogeography, 11, 377-392.

Ramier, D., N. Boulain, B. Cappelaere, F. Timouk, M. Rabanit, C. R. Lloyd, S. Boubkaroui, F. Descroix, and V. Wawrzyniak (2009), Towards an understanding of coupled physical and biological process in the cultivated Sahel–1. Energy and water, J. Hydrol., 375, 204–216, doi:10.1016/j.jhydrol.2008.12.002.

Randerson, J. T., et al. (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, Global Change Biol., doi:10.1111/j.1365-2486.2009.01912.x.

Reichler, T., and J. Kim (2008), How well do coupled models simulate today's climate?, Bull. Am. Meteorol. Soc., 89(3), 303–311, doi:10.1175/BAMS-89-3-303

Richards, J. F. (1984), Documenting environmental history: Global pattern of land conversion, Environment, 26(9), 6–13, 34–38

Roads et al., 2003, GCIP water and energy budget synthesis (WEBS). J. Geophys. Res., 108(D16), 8609,doi:10.1029/2002JD002583

Roads, J., and A. Betts (2000), NCEP-NCAR and ECMWF reanalysis surface water and energy budgets for Mississippi river basin, J. Hydrometeorol., 1, 88–94, doi:10.1175/1525-7541(2000)001<0088:NNAERS>2.0.CO;2.

Roads, J., et al. (2003), GCIP water and energy budget synthesis (WEBS), J. Geophys. Res., 108(D16), 8609, doi:10.1029/2002JD002583.

Roads, J., M. Kanamitsu, and R. Stewart (2002), CSE water and energy budgets in the NCEP-DOE Reanalysis II, J. Hydrometeorol., 3, 227–248, doi:10.1175/1525-7541(2002)003<0227:CWAEBI>2.0.CO;2.

Robock, A., et al. (2003), Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season, J. Geophys. Res., 108(D22), 8846, doi:10.1029/2002JD003245.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zaho, M. Baldwin, N. W. Junker, and Y. Lin (1996), Changes to the operational "early" Eta analysis/forecast system at the National Centers for Environmental Prediction, 1996, Weather Forecast., 11, 391–413, doi:10.1175/1520-0434 (1996)011<0391:CTTOEA>2.0.CO;2.

Rudel et al., 2005, Forest Transition: towards a global understanding of land use change, Global Environmental Change, 15, 23-31.

Running, S. W., D. D. Baldocchi, D. P. Turner, S. T. Gower, P. S. Bakwin, and K. A. Hibbard (1999), A Global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling, and EOS satellite data, Remote Sens. Environ., 70, 108–127, doi:10.1016/S0034-4257(99)00061-9.

Ryu, Y., D. D. Baldocchi, S. Ma, and T. Hehn (2008), Interannual variability of evapotranspiration and energy exchange over an annual grassland in California, J. Geophys. Res., 113, D09104, doi:10.1029/2007JD009263.

Saha, S., et al. (2010), The NCEP Climate Forecast System Reanalysis, Bull. American Meteorological Society, 1015-1057.

Santanello, J. A., Jr., C. D. Peters-Lidard, S. V. Kumar, C. Alonge, and W. K. Tao (2009), A modeling and observational framework for diagnosing local land-atmosphere coupling on diurnal time scales, J. Hydrometeorol., 10, 577–599, doi:10.1175/2009JHM1066.1.

Scanlon, B. R., I. Jolly, M. Sophocleous, and L. Zhang (2007), Global impacts of conversions from natural to agricultural ecosystems on water resources: Quantity versus quality, Water Resour. Res., 43, W03437, doi:10.1029/2006WR005486.

Schär, C., D. Lüthi, U. Beyerle, and E. Heise (1999), The soil-precipitation feedback: A process study with regional climate model, J. Clim., 12(3), 722–741, doi:10.1175/1520-0442(1999)012<0722:TSPFAP>2.0.CO;2.

Schott, C. A. (1881), Tables and results of the precipitation in rain and snow in the United States, and some stations in adjacent parts of North America, and in Central and South America, Smithson. Contrib. Knowl., 353, 112–182.

Schuol, J., K. C. Abbaspour, H. Yang, R. Srinivasan, and A. J. B. Zehnder (2008), Modeling blue and green water availability in Africa, Water Resour. Res., 44, W07406, doi:10.1029/2007WR006609.

Shelton, M. L. (2009), Hydroclimatology: perspective and applications, Cambridge University Press, UK.

Sohl, T. L., K. L. Sayler, M. A. Drummond, and T. R. Loveland (2007), The FORE-SCE mode: a practical approach for projecting land cover change using scenarios-based modeling, Journal of Land Use Science, 2, 2, 103-126.

Sohl, T. L., T. R. Loveland, B. M. Sleeter, K. L. Sayler, and C. A. Barnes (2010), Addressing foundational elements of regional land-use change forecasting, Landscape Ecol., 25, 233-247.

Sohl, T., and K. Sayler (2008), Using the FORE-SCE model to project land-cover change in the southeastern United States, Ecological Modeling, 219, 49-65.

Solomon, S. et al. (2007): Technical Summary. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the FourthAssessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis,K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY,USA.

Sridhar, V., R. L. Elliott, F. Chen, and J. A. Brotzge (2002), Validation of the NOAH-OSU land surface model using surface flux measurements in Oklahoma, J. Geophys. Res., 107(D20), 4418, doi:10.1029/2001JD001306.

Steyaert, L. T., and R. G. Knox (2008), Reconstructed historical land cover and biophysical parameters for studies of land-atmospheric interactions within the eastern United States, J. Geophys. Res., 113, D02101, doi:10.1029/2006JD008277.

Stöckli, R., D. M. Lawrence, G.-Y. Niu, K. W. Oleson, P. E. Thornton, Z.-L. Yang, G. B. Bonan, A. S. Denning, and S. W. Running (2008), Use of FLUXNET in the Community Land Model development, J. Geophys. Res., 113, G01025, doi:10.1029/2007JG000562.

Strack, J. E., R. A. Pielke Sr., L. T. Steyaert, and R. G. Knox (2008), Sensitivity of June near-surface temperatures and precipitation in the eastern United States to historical land cover changes since European settlement, Water Resour. Res., 44, W11401, doi:10.1029/2007WR006546.

Tanner, C. B., and E. R. Lemon (1962), Radiant energy utilized in evapotranspiration, Agron. J., 54, 207–212, doi:10.2134/agronj1962.00021962005400030008x.

United States Department of Agriculture (2011), Crop Production 2010 Summary.

Vinnikov, K. Y., A. Robock, N. A. Speranskaya, and C. A. Schlosser (1996), Scales of temporal and spatial variability of midlatitude soil moisture, J. Geophys. Res., 101(D3), 7163–7174, doi:10.1029/95JD02753.

Vorosmarty C. J., P. Green, J. Salisbury, and R. B. Lammers, 2000. Vulnerability from climate change and population growth. Sceince, 289, 284 – 288.

Vose, R. S., C. N. Williams, T. C. Peterson, T. R. Karl, and D. R. Easterling (2003), An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network, Geophys. Res. Lett., 30(20), 2046, doi:10.1029/2003GL018111.

Wahl, E. W. (1968), A comparison of the climate of the eastern United States during the 1830's with the current normals, Mon. Weather Rev., 96(2), 73–82, doi:10.1175/1520-0493(1968)96<73:ACOTCO>2.0.CO;2.

Waisanen, P. J., and N. B. Bliss, 2002, Changes in population and agricultural land in conterminous United States counties, 1790 to 1997, Global Biogeochem. Cycles, 16(4), 1137, doi:10.1029/2001GB001843.

Wang, A., K. Y. Li, and D. P. Lattenmaier (2008), Integration of the variable infiltration capacity model soil hydrology scheme into the community land model, J. Geophys. Res., 113, D09111, doi:10.1029/2007JD009246.

Werth, D., and R. Avissar (2002), The local and global effects of Amazon deforestation, J. Geophys. Res., 107(D20), 8087, doi:10.1029/2001JD000717.

Whitney, G. G. (1994), From Coastal Wilderness to Fruited Plain: A History of Environmental Change in Temperate North America from 1500 to the Present, Cambridge Univ. Press, New York.

Williams, M. (1989), American and Their Forest A Historical Geography,Cambridge Univ. Press, New York.

Wilson, K., et al. (2002), Energy balance closure at FLUXNET sites, Agric.For. Meteorol., 113, 223–243, doi:10.1016/S0168-1923(02)00109-0.

Yang G., L.C. Bowling, K. A. Cherkauer, B. C. Pijanowski, and D. Niyogi (2009): Hydroclimatic Response of Watersheds to Urban Intensity- An Observational and Modeling Based Analysis for the White River Basin, Indiana, Journal of Hydrometeorology, DOI: 10.1175/2009JHM1143.1

Yeager, S. G., C. A. Schields, W. G. Large, and J. J. Hack (2006), The low resolution CCSM3, J. Clim., 19, 2545–2566, doi:10.1175/JCLI3744.1.

Yeh, P. J. F., and J. S. Famiglietti (2009), Regional groundwater evapotranspiration in Illinois, J. Hydrometeorol., 10(2), 464–478, doi:10.1175/2008JHM1018.1.

Zeng, X., M. Shaikh, Y. Dai, R. E. Dickinson, and R. Myneni (2002), Coupling of the common land model to the NCAR community climate model, J. Clim., 15, 1832–1854, doi:10.1175/1520-0442(2002) 015<1832:COTCLM>2.0.CO;2.

Zhang, L., K. Hickel, W. R. Dawes, F. H. S. Chiew, A. W. Western, and P. R. Briggs (2004), A rational function approach for estimating mean annual evapotranspiration, Water Resour. Res., 40, W02502, doi:10.1029/2003WR002710.

Zhang, L., N. Potter, K. Hickel, Y. Zhang, and Q. Shao (2008), Water balance modeling over variable time scales based on the Budyko framework—Model development and testing, J. Hydrol., 360, 117–131, doi:10.1016/j.jhydrol.2008.07.021.

Zhang, L., W. R. Dawes, and G. R. Walker (2001), Response of mean annual evapotranspiration to vegetation changes at catchment scale, Water Resour. Res., 37(3), 701–708, doi:10.1029/2000WR900325.

Zhang, Y. K., and K. E. Schilling (2006), Increasing streamflow and base flow in Mississippi River Basin since 1940: Effect of land use change, J. Hydrol., 324, 412–422, doi:10.1016/j.jhydrol.2005.09.033.

Zhu, C., L. R. Leung, D. Gochis, Y. Qian, and D. P. Lettenmaier (2009), Evaluating the influence of antecedent soil moisture on variability of the North American monsoon precipitation in the coupled MM5/VIC modeling system, J. Adv. Model. Earth Syst., 1, 13, 22pp.

APPENDICES

Appendix A

Biophysical drivers of Land Cover Change
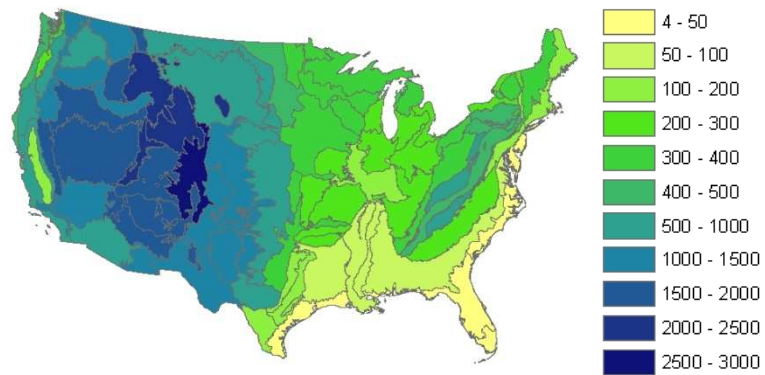
(a)   Elevation (meter)
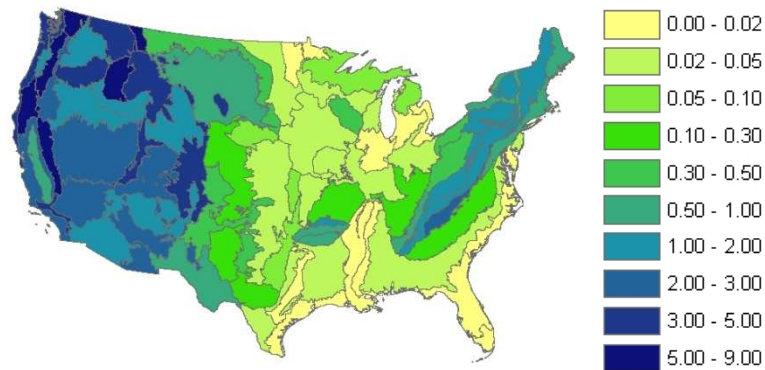
(b)   Slope (%)
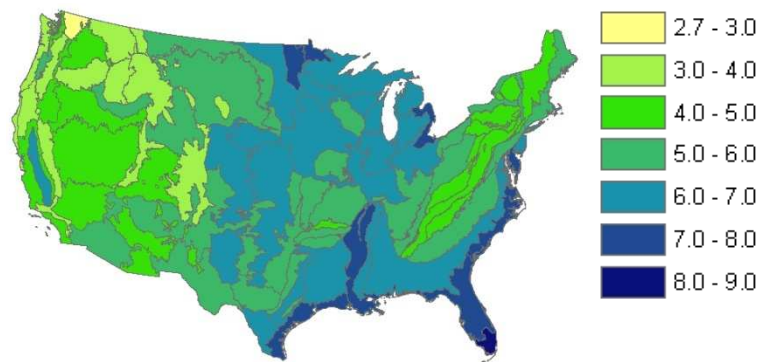
(c)   Topographic index
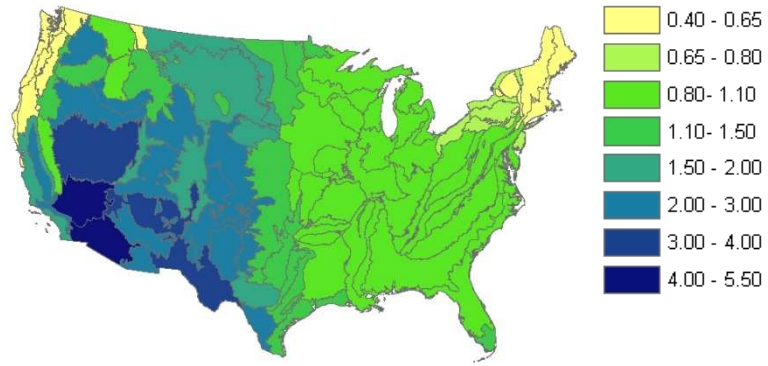
(d)   Dryness index
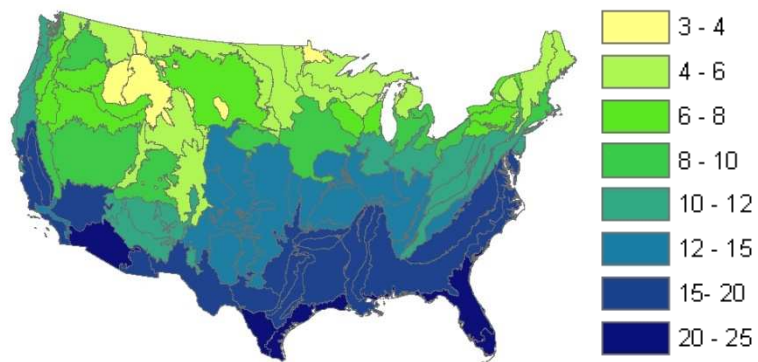
(e)   Annual temperature (^{O}C)
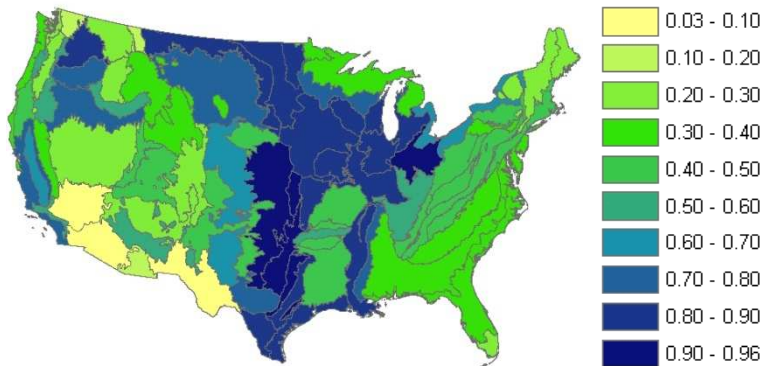
(f)   Crop suitability index
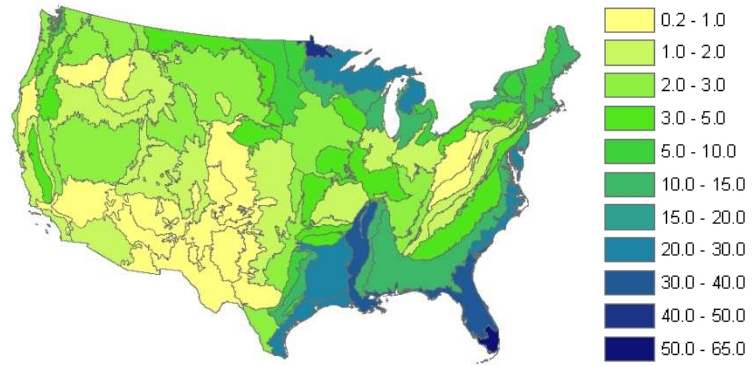
(g)   Water and wetland Percentage

Fig. A.1: Biophysical drivers of land-cover change

Urban Cover Model Results
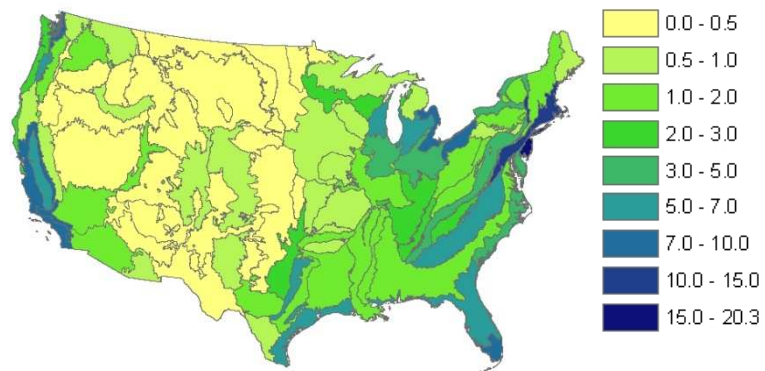
Fig. A.2: %urban cover increase
(2000-1900)

Population density for 1850, 1900, 1950, and 2000
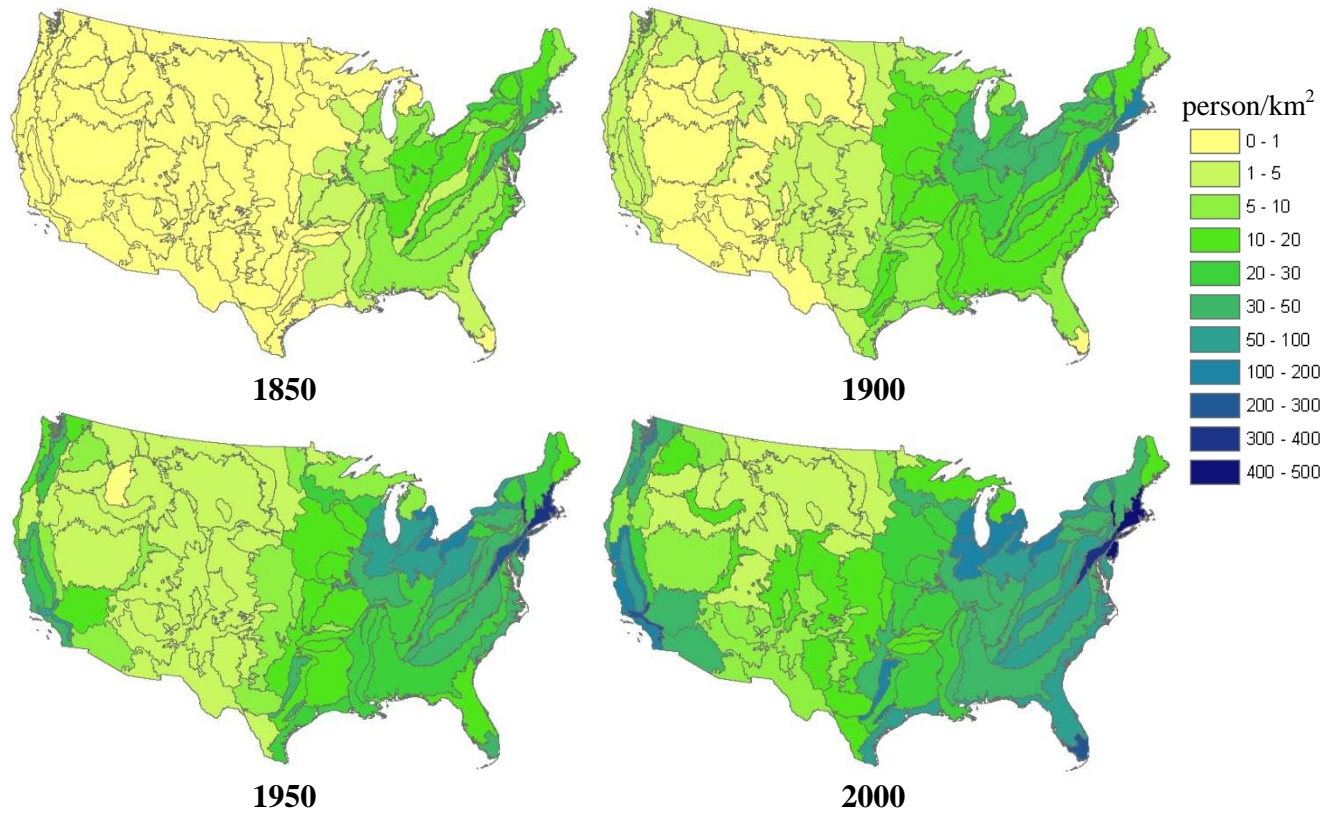
Fig. A.3: Ecoregion average population density (person/km^2) for 1850, 1900, 1950, and 2000

Appendix B

Gamma and Beta Functions for Cropland Trajectory

First a gamma probability distribution (Eq. B.1) is fitted to *NCP* time series for each ecoregion, considering *NCP* as frequency and decade (1 to 16 for 1850 to 2000) as bin value, and then a beta probability distribution function (Eq. B.2) is fitted to absolute value of residuals *(|NCP(x) – Γ(x)|)*.

$$\Gamma(x) = \frac{hv_1}{\overline{)\alpha} \cdot \delta}\left(\frac{x-\theta}{\delta}\right) \cdot \exp\left(-\left(\frac{x-\theta}{\delta}\right)\right) \qquad for\ x > \theta \qquad (B.1)$$

$$\beta(x) = hv_2 \frac{(x-\theta)^{\lambda-1}(\sigma+\theta-x)^{\omega-1}}{\left(\dfrac{\overline{)\lambda} \cdot \sqrt{\omega}}{\overline{)\lambda+\omega}}\right) \cdot \sigma^{(\lambda+\omega-1)}} \qquad for\ \theta < x < \theta+\sigma \qquad (B.2)$$

$$\overline{)z} = \int_0^\infty e^{-u} \cdot u^{z-1} du \qquad (B.3)$$

Where $\Gamma(x)$ and $\beta(x)$ are gamma and beta probability distribution function, respectively, x is the decade number (1 to 16 for 1850 to 2000), $\theta\,(=0)$ is a threshold parameter, $h\,(=1)$ is the width of histogram interval, v is the vertical scaling factor (v_1 = sum of *NCP*, v_2 = sum of absolute value of residual between *NCP* and $\Gamma(x)$), σ (= 17) is a scale parameter in beta function, and it is kept constant for all ecoregions. Parameters α (a shape parameter) and δ (a scale parameter) for gamma function, and λ (a shape parameter) and ω (another shape parameter) for beta functions are first determined in SAS using the *NCP* time series for the ecoregion, then these parameters are optimized for minimum sum of square residual between observation (*NCP*) and (gamma + beta) function values using excel solver tool. Goodness of fit is determined using Nash-Sutcliffe Efficiency coefficient (*NSE*, Eq. B.4).

$$NSE_i = 1 - \frac{\displaystyle\sum_{x=1}^{16}(NCP_i(x)-[\Gamma_i(x)+\beta_i(x)])^2}{\displaystyle\sum_{x=1}^{16}(NCP_i(x)-\overline{NCP_i})^2} \qquad (B.4)$$

Appendix C

Multiple linear regression model results details

Independent variables in the analysis are: PD, ELEV, TI, DI, AT, SUIT, WP and dependent variables is CP. SL and UP are not included in the analysis because: (1) SL is highly correlated with TI (correlation coefficient = 0.84), and (2) UP is highly correlated with PD (population density based model output). Fig. C.1 show temporal evolution number of ecoregion included in the analysis and corresponding fractional area of the conterminous United States. Except for TI, all other variables were not normally distributed. Power transformation was successful in bringing most variables (PD, WT, ELEV, DI, and AT) into normal distribution, confirmed by Shapiro-Wilk test and quantile-quantile plot. In all decades SUIT, and for some decades (1900, 1910, 1920, 1940) CP did not come into normal distribution. However, power transformation did bring SUIT and CP closer to normal. Power transformation is performed for each decade and each variable separately, because of increasing number of ecoregion included in the analysis. Model residuals for all decades are found normally distributed.

Effect of population threshold and power transformation of the major finding of this study (Fig. 4.2) is analyzed by performing similar analysis including all 84 ecoregions starting from 1850, and all non transformed variables. Major finding were similar as discussed in the main text (not shown).
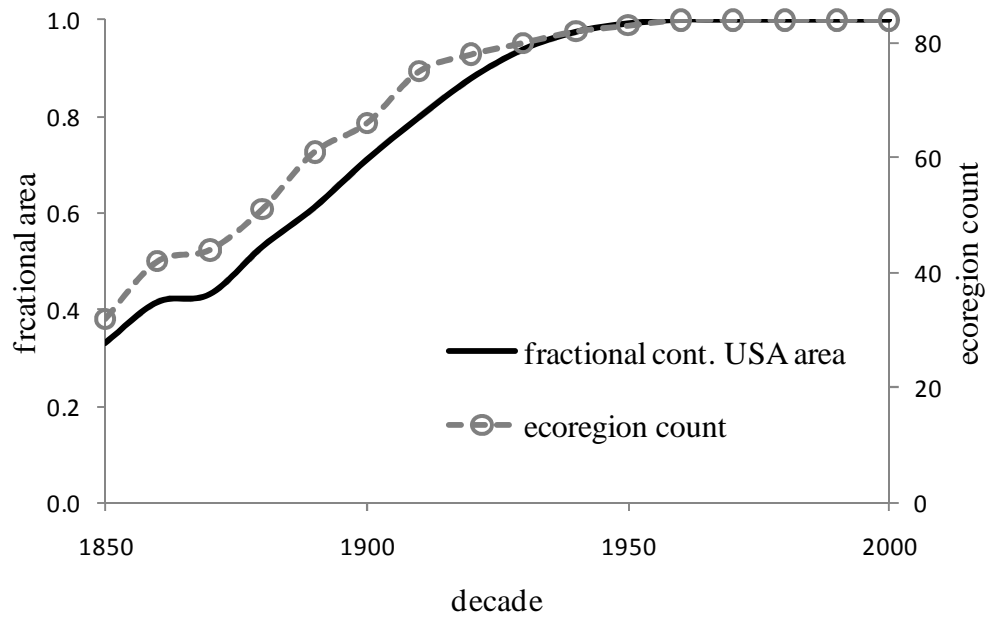
Fig. C.1: Number of ecoregions having population density greater than 1 person/km^2

Appendix D

Urban Cover Model

NLCD 2001 Urban Cover data (30 m resolution) is regrided to 0.5^O resolution (Fig. D.1). A regression tree (recursive partition) model is developed for %urban cover based on population density as independent variable. The best fit regression tree model (Fig. D.2) is obtained corresponding to minimum deviance in the data (Fig. D.3). The regression tree model is a step model for %urban cover (Fig. D.4). A continuous %urban cover model is obtained by joining midpoint of each interval (Fig. D.4), and this model is used for this study (Eq. D.1).

$$
\left.
\begin{aligned}
y_{i,t} &= 0.1482 * x_{i,t} + \varepsilon_{i,\,2000} & &\text{for } x_{i,t} < 14.92 \\
y_{i,t} &= 0.075 * x_{i,t} + 1.09 + \varepsilon_{i,\,2000} & &\text{for } 14.92 \le x_{i,t} < 56.13 \\
y_{i,t} &= 0.126 * x_{i,t} - 1.77 + \varepsilon_{i,\,2000} & &\text{for } 56.13 \le x_{i,t} < 93.83 \\
y_{i,t} &= 0.0673 * x_{i,t} + 3.74 + \varepsilon_{i,\,2000} & &\text{for } 93.83 \le x_{i,t} < 212.04 \\
y_{i,t} &= 0.0634 * x_{i,t} + 4.55 + \varepsilon_{i,\,2000} & &\text{for } 212.04 \le x_{i,t} < 361.86 \\
y_{i,t} &= 0.0572 * x_{i,t} + 6.81 + \varepsilon_{i,\,2000} & &\text{for } 361.86 \le x_{i,t} < 598.03 \\
y_{i,t} &= 0.0681 * x_{i,t} + 0.27 + \varepsilon_{i,\,2000} & &\text{for } 598.03 \le x_{i,t} < 906.38 \\
y_{i,t} &= 74 + \varepsilon_{i,\,2000} & &\text{for } x_{i,t} \ge 906.38
\end{aligned}
\right\}
\qquad \text{D.1}
$$

where $y_{i,t}$ is %urban cover, and $x_{i,t}$ is population density (person/km^2) for grid cell *i* and year *t*. $\varepsilon_{i,\,2000}$ is the error for grid cell *i* and year *2000*. The error ($\varepsilon_{i,\,2000}$) is obtained corresponding to NLCD 2001 urban cover data and population density for year 2000 at 0.5^O resolution, and this error is kept constant for all decades (1850 to 2000). The error could be due to coarse resolution population density data (county level, ~42km) in comparison to fine resolution (30m) NLCD2001 urban cover data. The model is run for 1850 to 1990 period based on population density data to obtain hindcast of %urban cover for the conterminous United States. Model performances for the regression tree model and the final model are shown in Table D.1.
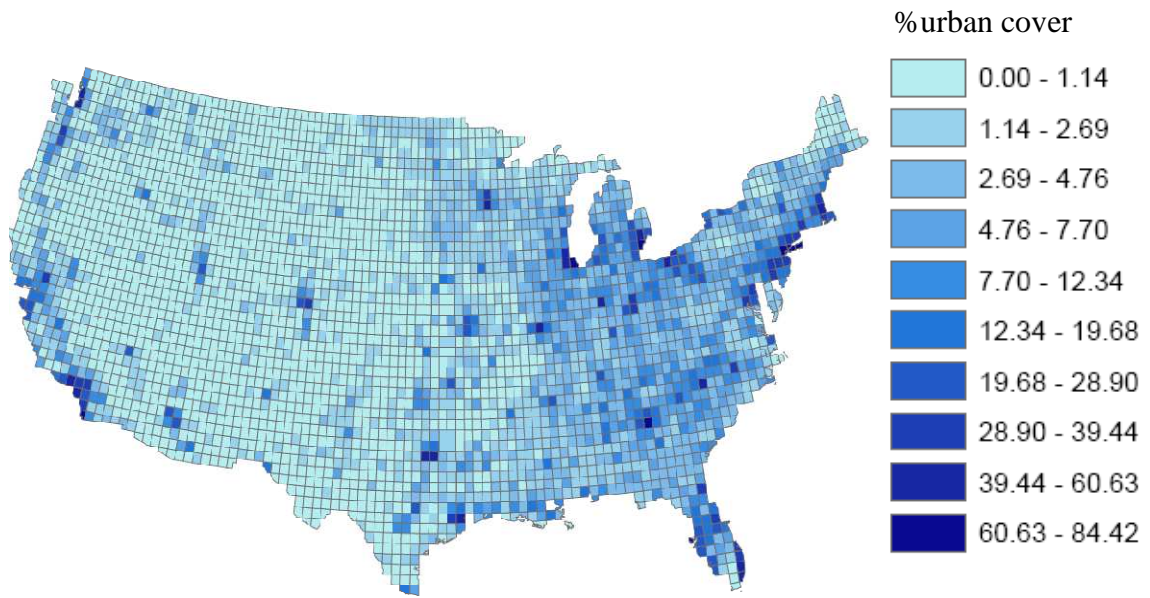
Fig. D.1: %urban cover for 2000 at 0.5$^{\text{O}}$ resolution
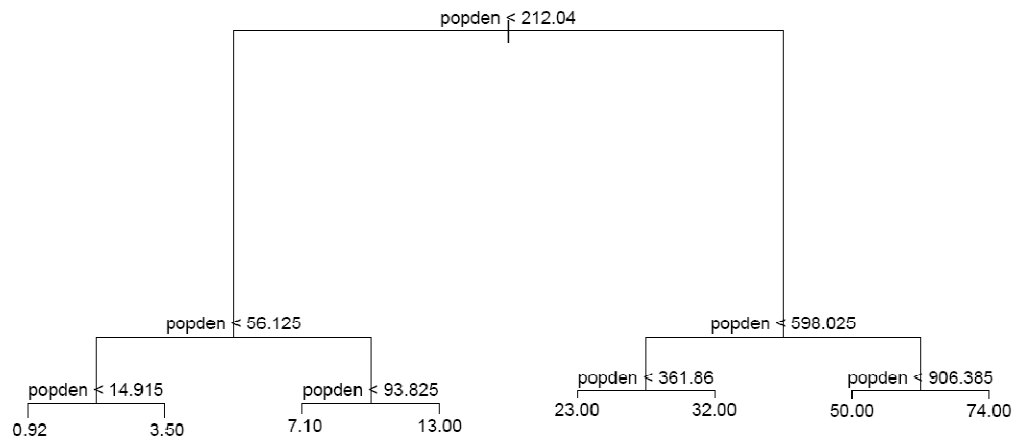
Fig. D.2: The Regression Tree model for %urban cover (popden: population density)
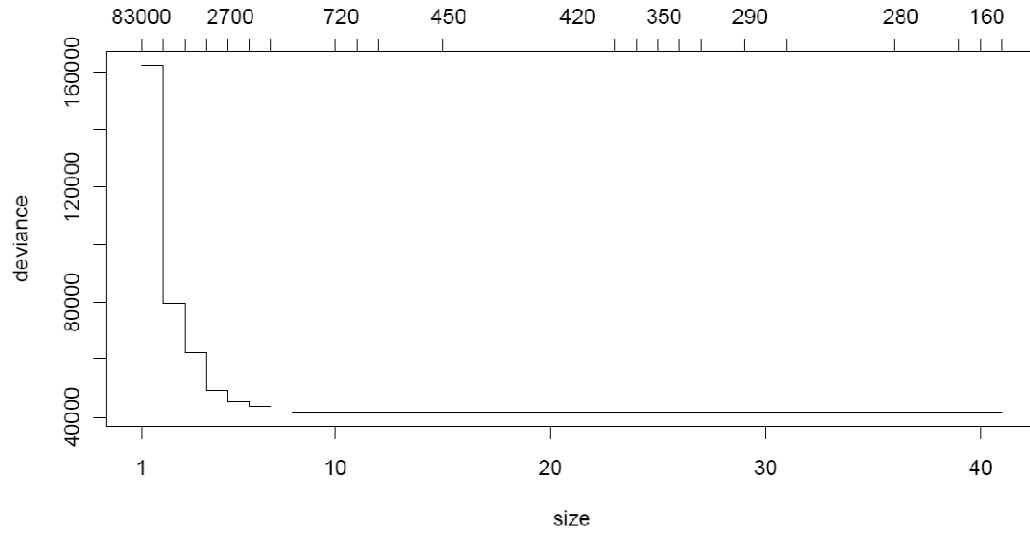
Fig. D.3: Tree size versus deviance plot

Table D.1: Urban cover model performance.

	model1	model2
RMSE	3.3	3.4
RSQ	0.78	0.77

Model 1 is the regression tree (step model), and model 2 is final continuous model used in this study. Although model performance is slightly deteriorated in the final model, because of continuity reason this model is selected for the study.
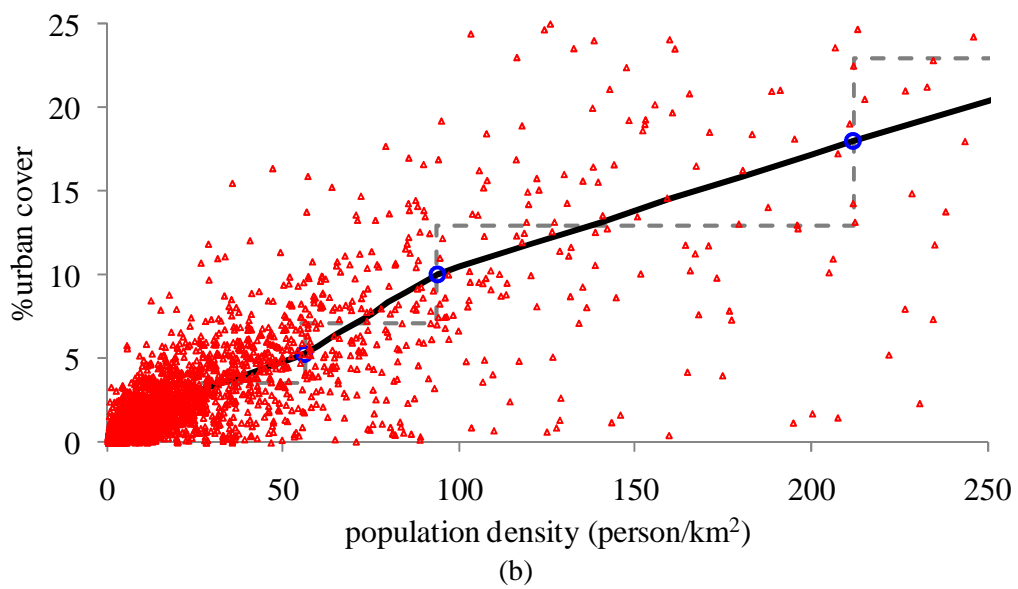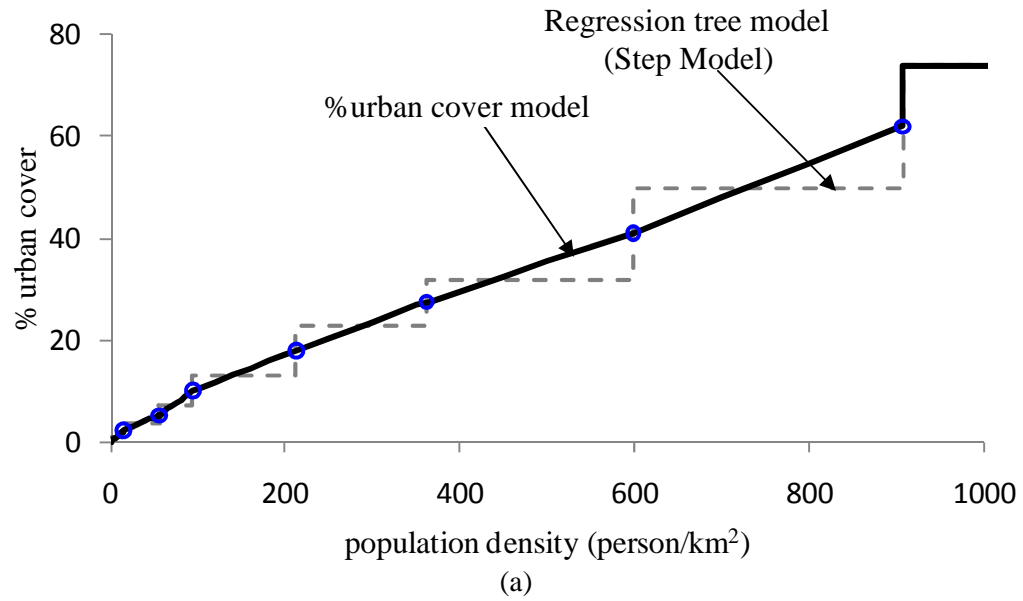
Fig. D.4: %urban cover model for the conterminous United States. (a) full model, (b) enlarged view near origin along with data (red triangles) for 2000

Appendix E
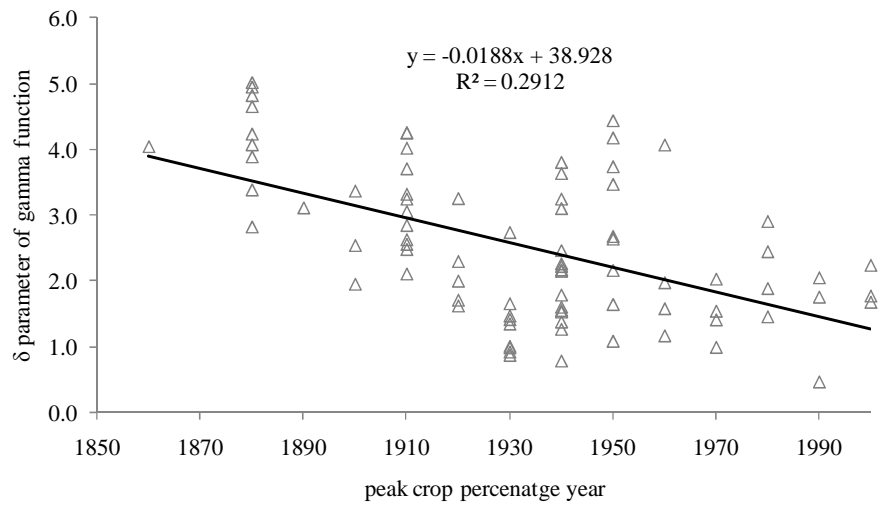
Fig. E.1: Scatter plot of δ parameter of gamma function with peak crop percentage year
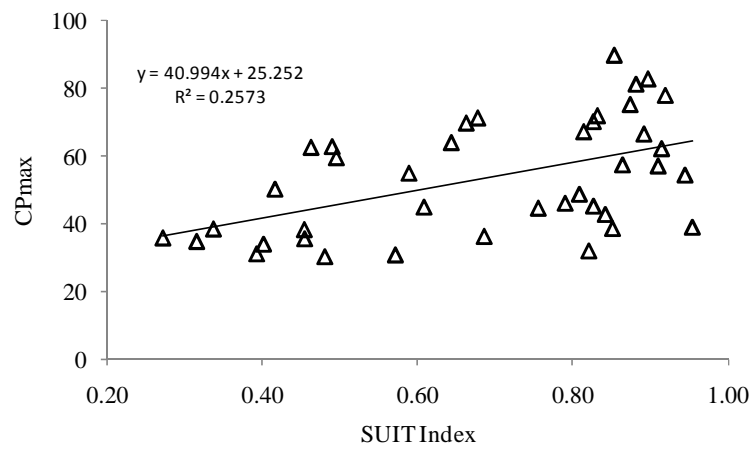
Fig. E.2: Scatter plot of SUIT index and CP_{max} for major crop region (historical) in USA

Table E.1: Gamma and Beta function parameters for all ecoregions. Note: ArB2G* [ArB2G$_{(1940-2000)}$].

eco#	Gamma		Beta Function			ν1	ν2	ν2/ν1	NSC	CPmax	Cpmax year	ArB2G*	SUIT
	α	δ	λ	ω	λ/ω								
eco1	3.38	2.82	1.75	0.24	7.33	1052	239	0.23	0.65	8.6	1880	0.27	0.36
eco2	7.05	1.64	0.82	0.70	1.18	879	123	0.14	0.95	6.2	1950	0.12	0.24
eco3	2.57	4.44	4.19	1.30	3.23	1234	279	0.23	0.92	18.9	1950	0.57	0.51
eco4	2.82	4.18	3.80	1.37	2.76	1130	269	0.24	0.91	8.3	1950	0.55	0.28
eco5	3.09	3.74	3.47	1.04	3.33	1171	277	0.24	0.93	10.3	1950	0.51	0.39
eco6	2.99	3.47	3.02	1.02	2.96	1204	234	0.19	0.90	20.3	1950	0.43	0.76
eco7	2.78	4.07	4.63	1.25	3.70	1226	295	0.24	0.93	36.3	1960	0.59	0.69
eco8	4.19	2.67	3.02	0.88	3.42	976	195	0.20	0.94	12.4	1950	0.36	0.51
eco9	4.89	2.64	3.44	1.49	2.31	903	213	0.24	0.97	8.2	1950	0.39	0.56
eco10	6.06	2.05	4.80	1.08	4.46	947	223	0.24	0.97	38.6	1990	0.39	0.85
eco11	5.79	2.03	6.93	1.25	5.57	1070	212	0.20	0.98	9.0	1970	0.36	0.71
eco12	7.43	1.88	4.52	1.51	2.99	742	225	0.30	0.97	18.3	1980	0.48	0.58
eco13	5.24	2.24	5.44	0.97	5.63	1047	197	0.19	0.98	2.2	2000	0.33	0.28
eco14	5.20	2.16	3.82	0.67	5.68	935	194	0.21	0.90	2.1	1950	0.32	0.07
eco15	7.62	1.54	8.08	1.38	5.87	984	173	0.18	0.99	7.3	1970	0.29	0.30
eco16	7.60	1.52	6.33	0.99	6.37	985	186	0.19	0.97	4.1	1940	0.28	0.25
eco17	8.00	1.46	9.71	1.36	7.15	958	166	0.17	0.99	8.1	1930	0.27	0.34
eco18	7.75	1.57	7.46	1.28	5.83	917	159	0.17	0.99	3.3	1960	0.27	0.36
eco19	5.88	1.97	6.78	1.06	6.37	1047	194	0.19	0.98	6.2	1960	0.31	0.48
eco20	7.47	1.67	6.76	1.16	5.85	891	183	0.21	0.99	2.9	2000	0.32	0.40
eco21	7.86	1.41	6.52	0.87	7.53	841	110	0.13	0.98	6.6	1930	0.19	0.24

Table E.1: Gamma and Beta function parameters for all ecoregions (continued)

eco#	Gamma		Beta Function			v1	v2	v2/v1	NSC	CPmax	Cpmax year	ArB2G*	SUIT
	α	δ	λ	ω	λ/ω								
eco22	9.03	1.26	10.05	0.71	14.23	753	112	0.15	0.96	1.9	1940	0.15	0.22
eco23	14.37	0.78	7.91	0.23	34.58	545	99	0.18	0.91	1.7	1940	0.07	0.53
eco24	13.79	0.86	0.50	0.10	4.88	634	170	0.27	0.83	2.0	1930	0.08	0.09
eco25	12.20	1.00	7.57	0.75	10.05	817	168	0.21	0.96	45.0	1930	0.21	0.61
eco26	12.32	0.91	7.19	0.44	16.22	672	138	0.21	0.93	26.1	1930	0.15	0.48
eco27	6.09	1.78	9.17	1.42	6.46	996	169	0.17	0.98	62.2	1940	0.32	0.91
eco28	4.92	1.95	5.07	0.50	10.15	935	196	0.21	0.89	57.1	1900	0.31	0.91
eco29	6.48	1.55	8.46	0.73	11.56	894	151	0.17	0.96	39.0	1940	0.24	0.95
eco30	4.17	2.74	3.55	0.67	5.31	1043	260	0.25	0.88	10.0	1930	0.42	0.70
eco31	5.14	3.10	1.75	0.97	1.80	861	368	0.43	0.78	11.8	1940	0.61	0.81
eco32	4.77	2.14	3.38	0.69	4.92	994	133	0.13	0.96	54.4	1940	0.23	0.94
eco33	4.72	2.22	4.63	0.59	7.87	952	174	0.18	0.93	42.8	1940	0.28	0.84
eco34	5.80	2.46	1.85	0.95	1.95	948	294	0.31	0.95	32.1	1940	0.40	0.82
eco35	4.72	2.16	2.67	0.32	8.26	936	148	0.16	0.87	24.4	1940	0.18	0.48
eco36	6.10	1.70	8.06	0.61	13.19	884	156	0.18	0.93	18.4	1920	0.22	0.44
eco37	6.53	1.62	8.69	0.78	11.13	918	150	0.16	0.95	30.8	1920	0.23	0.57
eco38	5.17	1.99	8.04	0.86	9.37	998	149	0.15	0.96	27.5	1920	0.25	0.44
eco39	4.53	2.29	9.42	1.41	6.70	1064	151	0.14	0.97	35.6	1920	0.31	0.45
eco40	3.94	2.48	10.37	1.71	6.07	1095	150	0.14	0.97	71.9	1910	0.35	0.83
eco41	7.70	1.75	4.08	0.99	4.13	795	221	0.28	0.97	11.2	1990	0.40	0.10
eco42	8.84	1.40	5.05	0.87	5.79	887	188	0.21	0.98	45.2	1970	0.28	0.83

Table E.1: Gamma and Beta function parameters for all ecoregions (continued)

eco#	Gamma		Beta Function			v1	v2	v2/v1	NSC	CPmax	Cpmax year	ArB2G*	SUIT
	α	δ	λ	ω	λ/ω								
eco43	11.91	0.98	9.60	0.69	13.94	713	171	0.24	0.93	24.9	1930	0.21	0.79
eco44	8.10	1.34	6.74	0.74	9.09	874	145	0.17	0.96	30.3	1930	0.22	0.48
eco45	3.36	2.84	1.71	16.08	0.11	1122	187	0.17	0.91	38.5	1910	0.00	0.34
eco46	6.88	1.65	8.44	1.37	6.17	1038	184	0.18	0.98	75.2	1930	0.31	0.87
eco47	4.05	2.63	8.13	1.52	5.34	1202	214	0.18	0.98	81.2	1910	0.41	0.88
eco48	7.22	1.60	6.06	0.82	7.34	878	174	0.20	0.95	89.7	1940	0.28	0.85
eco49	8.94	1.45	7.83	1.40	5.60	793	169	0.21	0.99	12.7	1980	0.32	0.35
eco50	8.09	1.37	11.60	1.52	7.65	856	98	0.11	0.99	12.0	1940	0.19	0.38
eco51	4.98	2.26	6.15	1.35	4.57	1015	160	0.16	0.99	46.1	1940	0.31	0.79
eco52	3.17	3.24	7.65	1.62	4.71	1255	214	0.17	0.98	57.4	1910	0.45	0.86
eco53	2.40	4.23	5.56	1.57	3.54	1327	223	0.17	0.97	66.5	1880	0.49	0.89
eco54	2.46	4.25	6.11	1.46	4.18	1350	249	0.18	0.97	82.7	1910	0.53	0.90
eco55	2.37	4.26	4.62	1.20	3.85	1284	225	0.18	0.90	77.9	1910	0.49	0.92
eco56	2.98	3.32	4.40	1.12	3.94	1158	162	0.14	0.97	67.2	1910	0.35	0.81
eco57	3.28	3.25	4.43	1.10	4.01	1179	200	0.17	0.97	69.7	1920	0.39	0.66
eco58	1.79	4.07	1.16	6.14	0.19	833	81	0.10	0.98	35.9	1880	0.00	0.27
eco59	1.58	4.04	0.83	4.60	0.18	794	76	0.10	0.99	50.3	1860	0.01	0.42
eco60	2.05	3.89	0.64	0.57	1.14	1026	97	0.09	0.96	62.8	1880	0.16	0.49
eco61	1.91	4.65	0.50	0.47	1.05	1057	133	0.13	0.95	71.2	1880	0.18	0.68
eco62	2.43	3.37	0.50	0.47	1.05	1029	96	0.09	0.96	26.9	1900	0.13	0.35

Table E.1: Gamma and Beta function parameters for all ecoregions (continued)

eco#	Gamma		Beta Function			v1	v2	v2/v1	NSC	CPmax	Cpmax year	ArB2G*	SUIT
	α	δ	λ	ω	λ/ω								
eco64	1.75	5.02	0.70	0.70	0.99	1179	154	0.13	0.94	62.5	1880	0.21	0.46
eco65	3.45	3.64	0.84	2.29	0.37	1152	280	0.24	0.85	31.2	1940	0.08	0.39
eco66	3.69	2.55	0.89	8.32	0.11	1021	146	0.14	0.92	34.0	1910	0.00	0.40
eco67	3.23	3.06	0.98	7.06	0.14	1138	185	0.16	0.92	38.3	1910	0.00	0.45
eco68	3.64	3.25	0.80	1.84	0.43	1099	222	0.20	0.91	28.4	1940	0.09	0.43
eco69	4.09	2.10	0.61	8.43	0.07	929	128	0.14	0.94	28.4	1910	0.00	0.47
eco70	3.25	2.54	0.63	12.01	0.05	961	123	0.13	0.94	59.5	1900	0.00	0.49
eco71	3.47	4.02	1.42	3.66	0.39	1253	389	0.31	0.92	54.9	1910	0.07	0.59
eco72	2.74	3.71	7.09	1.48	4.78	1235	195	0.16	0.98	70.2	1910	0.44	0.83
eco73	5.23	2.90	1.29	0.71	1.82	949	349	0.37	0.96	48.7	1980	0.46	0.81
eco74	3.82	3.80	1.02	1.70	0.60	1164	391	0.34	0.88	44.6	1940	0.22	0.76
eco75	6.26	2.44	1.06	0.96	1.10	980	366	0.37	0.98	10.9	1980	0.36	0.35
eco76	33.59	0.46	3.58	1.15	3.11	489	157	0.32	0.99	14.5	1990	0.37	0.26
eco77	10.62	1.08	8.19	0.85	9.68	888	127	0.14	0.98	6.3	1950	0.17	0.14
eco78	3.20	3.11	1.37	0.27	5.05	992	238	0.24	0.72	6.0	1890	0.30	0.44
eco79	13.92	0.99	2.63	1.67	1.58	625	161	0.26	0.88	3.1	1970	0.23	0.17
eco80	7.15	1.77	6.21	1.09	5.71	857	189	0.22	0.97	6.7	2000	0.34	0.75
eco81	10.58	1.16	10.42	2.34	4.46	771	158	0.20	0.97	6.4	1960	0.32	0.03
eco82	2.19	3.38	0.46	3.02	0.15	815	78	0.10	0.96	18.1	1880	0.01	0.23
eco83	1.76	4.95	0.98	0.90	1.08	1059	136	0.13	0.98	63.9	1880	0.23	0.64
eco84	1.72	4.82	0.67	1.07	0.63	1035	119	0.11	0.97	34.8	1880	0.15	0.31

Table E.2: Analysis of rising limb in 17 major crop ecoregions

Sl No.	Year (CPmax)	eco#	T1	F (T1)	T2	F (T2)	T1-T2 (MRD)
1	1880	eco53	1	27	6	101	5
2	1900	eco28	3	26	8	95	5
3	1910	eco55	1	27	6	99	5
4	1910	eco47	3	34	8	103	5
5	1910	eco52	2	31	7	102	5
6	1920	eco39	3	27	8	96	5
7	1930	eco25	7	33	11	99	4
8	1930	eco44	5	29	10	97	5
9	1940	eco29	4	29	9	96	5
10	1940	eco27	4	26	9	98	5
11	1940	eco48	5	27	10	90	5
12	1940	eco34	5	29	13	93	8
13	1940	eco51	4	32	10	92	6
14	1960	eco7	2	32	15	97	13
15	1970	eco42	6	28	12	98	6
16	1980	eco73	5	31	16	93	11
17	1990	eco10	5	28	12	94	7

T1: First decade when rising limb falls between 25 and 35, T2: Decade when rising limb reaches its maximum value. F(x) is the gamma + beta function for given x. Average F(T1) = 29, and average F(T2) = 97.
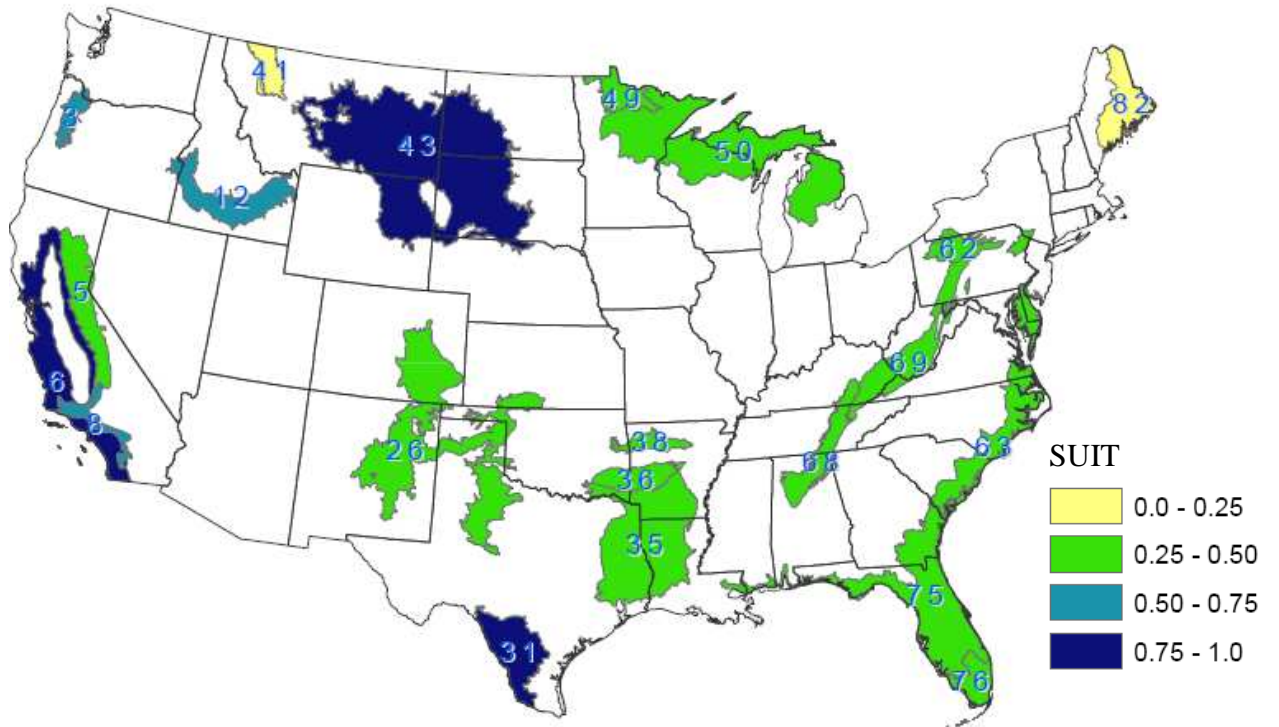
Fig. E.3: Ecoregions with $10\% \leq CP_{max} < 30\%$.

VITA

VITA

Sanjiv Kumar was born in Bihar, India in 1978. He received his Bachelor of Technology degree in Civil Engineering from Banaras Hindu University, Varanasi, India in 2001. He joined the graduate program of Purdue University in January 2007 and received his Master of Science in Civil Engineering and Doctor of Philosophy degree in May 2008, and May 2011, respectively. He worked for five years in a power plant equipment manufacturing and construction company before joining the graduate program.