The Dissertation Committee for Xitian Cai Certifies that this is the approved
version of the following dissertation:

Hydrological Assessment and Biogeochemical Advancement of the Noah-MP Land Surface Model

Committee:

Zong-Liang Yang, Supervisor

Robert E. Dickinson

Bridget R. Scanlon

Joshua B. Fisher

Daniel O. Breecker

Hydrological Assessment and Biogeochemical Advancement of the Noah-MP Land Surface Model

by

Xitian Cai, B.E.; M.E.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2015

Dedication

To my parents, who give me the freedom to follow my heart.

To my wife, whose understanding and support make this possible.

Acknowledgements

Preface

Being trained as a land surface modeler, I have been working on various topics of land surface modeling during my Ph.D. studies, with the ultimate goal of improving Noah-MP not only for climate modeling but also for application-based studies. When I first entered the Department of Geological Sciences, it was at the time that Noah-MP had just been developed and evaluated at local scale and a few global river basins. The measurements for local scale evaluation only included about two months of surface fluxes in the First ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment (FIFE) domain (15 km × 15 km). The evaluation for global river basins used Illinois soil moisture, GRACE TWS, GRDC runoff, and Canadian Meteorology Center snow depth and SWE data sets, which may not be considered the "best" available observations from the hydrological perspective. Clearly, we needed more comprehensive evaluation for Noah-MP. Therefore, in Chapter 2, I evaluated Noah-MP's performance in simulating the hydrological cycle for the Mississippi River Basin. I used measurements that are considered high quality for this region such as streamflow and groundwater measurements from USGS, ET measurement from AmeriFlux, and soil moisture from SCAN. The results from this chapter are encouraging. The simulated runoff and soil moisture are consistent with USGS and SCAN measurements, respectively. The model can reproduce the AmeriFlux ET for natural lands. Some disagreements were also found such as the much shallower simulated water table depth and the early fast growing ET over croplands. By doing the work presented in Chapter 2, I learned how to use Linux systems, Vim text editor, NCAR Command Language, and most importantly Noah-MP.

I presented my preliminary results from the Mississippi River Basin study at AGU Fall Meeting 2011. During the meeting, I had the opportunity to see other similar studies.

One thing that amazed me was the very high-quality streamflow data for 961 small river basins within CONUS. Fortunately, Dr. Youlong Xia from NOAA/NCEP came to my poster session and showed me that NCEP is the owner of the streamflow data for the 961 small river basins along with many other observational data sets and tools, which in combination make the so called "NLDAS test bed". Later, we decided that we should work together to evaluate Noah-MP within the NLDAS framework. Since Noah-MP is designed as NCEP's next generation LSM, it was clear that we should compare it with the current NCEP land model (i.e., Noah) and other major LSMs in the U.S. (i.e., VIC and CLM). This lead to the work presented in Chapter 3, in which we found no single model is superior to other models in simulating TWS, streamflow, soil moisture, and ET at the same time. Instead, each model has advantages in simulating one or two of these variables and has disadvantages in simulating others. Regarding Noah-MP, it shows the best performance in simulating soil moisture and among the best in simulating TWS. However, its simulated ET increases too fast during each growing season, which is associated with LAI simulation. The work in this chapter really gave me a clearer picture of the current performances of LSMs.

Knowing how difficult it is for Noah-MP to capture the seasonality of LAI, we wanted to improve this by considering the N stress on vegetation growth dynamically, which was not considered in the original model development. In addition, supported by the NASA Interdisciplinary Science Program, we proposed to enable Noah-MP's capability in simulating the movement of nutrients (such as N) from land to estuaries. From literature, we have seen that SWAT performed well in simulating the N cycle. So I planned to employ the parameterization of N dynamics from SWAT. Due to the partnership between NASA Jet Propulsion Laboratory (JPL) and UT Austin, I had the opportunity to interact with Dr. Joshua Fisher from JPL and noticed that he developed a plant N model called Fixation and

Uptake of N (FUN). Therefore, the N dynamic model integrated into Noah-MP includes two parts: the plant N model from FUN and the soil N model from SWAT. This is included in Chapter 4, from which we can see that the N model can generally reproduce the N state/flux variables at the evaluation site. In addition, the delayed peak LAI due to the inclusion of N dynamics is expected to improve the timing of ET modeling. Although our model cannot fully simulate the movement of N from land to estuaries due to the lack of a river N transport module, this is a big step towards that goal. The work in this chapter is my real contribution to the development of Noah-MP.

Making contributions to building a better model is really exciting for a land surface modeler like me. In the future, I hope to use land surface models as a powerful tool to answer some scientific questions for a better understanding of our mother Earth and to address our societal needs.

Hydrological Assessment and Biogeochemical Advancement of the Noah-MP Land Surface Model

Xitian Cai, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Zong-Liang Yang

Abstract

Land surface models (LSMs) simulate the energy, momentum, water, and carbon balance of the soil–vegetation–atmosphere system. As a key component of weather and climate models, LSMs play an important role in weather prediction and climate projections. Rapid growth in LSM development has resulted in both the improvement of existing process representation and the addition of new processes and functionalities. However, it is a challenge to evaluate the accuracy of energy, water, and nutrient fluxes simulated by LSMs, due to the lack of observational data and the complexity of interactions and feedback among different processes. Additionally, climate and terrestrial biosphere models consider nitrogen an important factor in limiting plant carbon uptake, while operational environmental models view nitrogen as the leading nutrient for causing eutrophication in water bodies. However, few LSMs include nitrogen dynamics and nitrogen leaching is usually not well parameterized; hence these LSMs are not feasible for applications-based modeling, particularly for land management and agricultural impacts.

Therefore, this dissertation uses the Noah-MP LSM to study the following three questions. (1) How do recent developments in Noah-MP improve its performance in hydrological modeling, based on a case study for the Mississippi River Basin? (2) Compared to other similar LSMs, what are the advantages and disadvantages of Noah-MP in assessing the water balance over the conterminous U.S.? (3) After coupling the Fixation and Uptake of Nitrogen plant model and the Soil and Water Assessment Tool soil nitrogen dynamics into Noah-MP, can this coupled model characterize the major nitrogen fluxes and how the nitrogen dynamics affect the carbon and water simulations? The main scientific findings are as follows. (1) Noah-MP shows significant improvement in modeling the major hydrological variables such as runoff, groundwater, evapotranspiration, soil moisture, and terrestrial water storage (TWS), which is very likely due to the incorporation of some major improvements into Noah-MP, particularly an unconfined aquifer storage layer for groundwater dynamics and an interactive vegetation canopy for dynamic leaf phenology. (2) Compared to other three LSMs, Noah-MP provides the best performance in simulating soil moisture and is among the best in simulating TWS. (3) The new Noah-MP with nitrogen dynamics performs well in capturing the major nitrogen state/flux variables (e.g., soil nitrate and nitrate leaching). (4) The addition of nitrogen dynamics in Noah-MP improves the modeling of the carbon and water cycles (e.g., net primary productivity and evapotranspiration).

Table of Contents

List of Tables

xiv

List of Figures

Chapter 1: Introduction

1.1 BACKGROUND

Land covers approximately 29.2% of the Earth's surface. Although this is a small fraction compared to ocean surface, land surface largely contributes to the spatial and temporal variability in our weather and climate systems, due primarily to the high spatial heterogeneity of soil, vegetation, and topography. Land areas not only host the bulk of the living biosphere (Verstraete, 1989) but also support human activities such as habitat, agriculture, and industry. These intensifying human activities have been shaping the formation of the human age—the Anthropocene (Monastersky, 2015).

Land surface processes generally refer to the exchanges of heat, water, and CO_2 fluxes among the following land components: soil, vegetation, snow, glaciers, inland water, animals, and much more (Yang, 2004). As a key component of weather and climate models, land surface models (LSMs) simulate the energy, momentum, water, and carbon balance of the soil–vegetation–atmosphere system (Foley et al., 1996). They play an important role in weather prediction and climate projections (Pitman, 2003).

Land surface processes that need to be modeled in weather and climate models are: energy and water budgets of the land surface, vertical heat and water transports of substrates, water transport associated with vegetation, exchange of heat and water vapor between the land surface and the atmosphere, and horizontal water transport at and below the surface (Warner, 2011). Some of these processes were well represented in early LSMs. However, some processes are still not well parameterized in the current generation of LSMs. For example, groundwater–surface water interaction—part of the vertical water transport—is recognized as important for climate modeling, particularly for regions where water table depth fluctuates within five meters of the land surface (Kollet and Maxwell,

2008). However, the simulations of water table depth by LSMs are still not very realistic, compared to *in situ* observations (Zampieri et al., 2012; Cai et al., 2014a).

This chapter provides some background information and a literature review about LSM development and evaluation, new available observational data sets, and the importance of the biogeochemical nitrogen modeling in supporting climate and environmental predictions. Finally, this chapter concludes with the research objective and an outline of this dissertation's organization.

1.2 REVIEW OF RELATED TOPICS

1.2.1 Land Surface Model Development and Evaluation

LSM development started with the "bucket" model (Manabe, 1969), which assumed soil moisture storage was a reservoir that was filled by rainfall and emptied by runoff. From then on, multilayer soils, vegetation canopy (Dickinson et al., 1986; Sellers et al., 1986), and the carbon cycle (Bonan, 1995; Sellers et al., 1996) were gradually included into LSMs. Rapid growth in LSM development has resulted in both the improvement of existing process representation and the addition of new processes and functionalities. During the past decade, widely adopted new model components include dynamic leaf/vegetation, groundwater, multilayer snow, biogeochemistry, irrigation, urban canopy, and much more.

Dynamic leaf/vegetation models were the development focus of various research groups in the world (Foley et al., 1996; Dickinson et al., 1998; Bachelet et al., 2001; Sitch et al., 2003). Although the introduction of dynamic leaf/vegetation models may sometimes degrade model performance, it provides a more feasible tool to investigate the impact of climate change, land cover and land use change, and urbanization on the climate and water and nutrient cycles (Lawrence et al., 2011), which allows further investigation of mitigation

2

measurements for global change. Another important improvement is the modeling of groundwater dynamics, which influences soil moisture and runoff generation and hence the interaction between the atmosphere and land. Several studies have incorporated groundwater models in LSMs with some being coupled with climate models (Fan et al., 2007; Miguez-Macho et al., 2007; Niu et al., 2007; Lo et al., 2010; Leung et al., 2011; Barlage et al., 2015).

Model evaluation helps the scientific community to identify good models or directions for future development. This also helps operational modeling centers such as the NOAA/NCEP (National Oceanic and Atmospheric Administration/National Centers for Environmental Prediction) implement new features to their next generation of weather and climate forecasting systems to improve their operational forecast capability. Besides weather and climate forecasts, LSMs have been widely used for drought monitoring, climate change impact assessment, and land cover/land use change assessment; therefore, the identification of a physically sound, appropriate complexity, computationally affordable, and easy to use model is critical to these applications.

1.2.2 New Observational Data

Because model development and evaluation are highly dependent on data availability, major efforts are being made in producing new and better observational data in the land community. In terms of evaluating LSM runoff, the difficulty is how to compare station gaged streamflow with grid based runoff, as simulated by models. Besides the UNH-GRDC gridded runoff, which has been used for LSM evaluation in previous years, the USGS developed a hydrologic-unit runoff (Brakebill et al., 2011) in 2008 and has updated it annually (personal communication with David M. Wolock at USGS). Although it is not grid based, it retains the accuracy from ground based measurements for each

hydrologic unit so that it better represents the spatial pattern. Groundwater evaluation has the same issue; however, Fan and Miguez-Macho (2011) have created a climatology water table depth map for the U.S., which contains 567,946 water table observational sites and is dense enough to show the groundwater spatial pattern for most parts of the U.S..

Compared to runoff evaluation, evapotranspiration (ET) evaluation is even more constrained by the lack of available data. Over recent decades, the global network of micrometeorological tower sites that coordinate eddy covariance measurements of CO_2, water vapor, and energy, FLUXNET (http://fluxnet.ornl.gov/) (Running et al., 1999; Baldocchi et al., 2001), has provided the most reliable ET measurement and has been considered valuable data for LSM development (Stöckli et al., 2008) and evaluation (Kumar and Merwade, 2011; Li et al., 2011). As part of FLUXNET, AmeriFlux features a much denser network of tower sites in the U.S. than other regional networks.

In addition, Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage (TWS) data (Tapley et al., 2004) can now validate the performance of LSMs in TWS simulation, which is an overall indicator assessing model proficiency in simulating water budget. During the past decade, Swenson and Milly (2006) pioneered the use of GRACE data in climate model evaluation, Niu et al. (2007) and Lo et al. (2010) successfully used GRACE data for the development of groundwater dynamics in LSMs, and van Dijk et al. (2011) have used GRACE data to evaluate the Australian Water Resources Assessment system and recommended necessary improvements in the model, such as better precipitation forcing and augmentation of groundwater dynamics. More information about applying GRACE data in model development and evaluation can be found in a review study by Güntner (2008).

1.2.3 Nitrogen Model for Assessing the Carbon–Nitrogen Feedback

The effect of N-limitation on plant growth has long been recognized; however, its effect on climate–carbon feedback has not been included in climate models until recently (Hungate et al., 2003). Researchers often report that the terrestrial biosphere can uptake 22 to 57% of expected anthropogenic CO_2 emissions that are predicted by an intermediate emissions scenario (Cramer et al., 2001; Houghton et al., 2001). Even without considering N-limitation, however, the Hadley Center ocean–atmosphere model version 3 (HadCM3) predicted the terrestrial biosphere would switch from a weak carbon sink to a strong carbon source by 2050, because the soil respiration rate will increase with temperature while the direct effect of CO_2 on photosynthesis will saturate (Cox et al., 2000). Kicklighter et al. (1999) compared four terrestrial biosphere models and demonstrated that models with nutrient cycling estimated much less CO_2 uptake than models without nutrient cycling. A field experiment study by Luo et al. (2004), which proposed that soil nitrogen availability will decrease with time when soil nitrogen is sequestrated into long-lived plant biomass and soil organic matter pools under elevated CO_2, attracted attention from the climate community. Since that time, several studies have explicitly incorporated nitrogen dynamics in climate models or in offline LSMs or dynamic global vegetation models (Thornton et al., 2007; Sokolov et al., 2008; Ostle et al., 2009; Thornton et al., 2009; Bonan and Levis, 2010; Fisher et al., 2010). Thornton et al. (2007) coupled the Community Land Model version 3 (CLM3) (Dickinson et al., 2006) with carbon and nitrogen dynamics of the terrestrial biogeochemistry model Biome-BGC version 4.1.2 (Thornton and Rosenbloom, 2005), in which the authors conclude that a model with C and N dynamics predicts 74% less of global terrestrial carbon uptake than the C-only model in response to increasing atmospheric CO_2 concentration. Fisher et al. (2010) provide similar conclusion but with a smaller percentage of reduction. According to Sokolov et al. (2008), who use the coupled

MIT Integrated Global Systems Model with carbon and nitrogen dynamics of Terrestrial Ecosystems Model, carbon uptake is also reduced in response to increasing atmospheric CO_2 concentration. In addition, they concluded that, in response to temperature increase, the terrestrial biosphere switches from being C-source to C-sink. Thornton et al. (2009) reach a similar conclusion from their coupled study using the Community Climate System Model (Collins et al., 2006) with carbon-only and carbon-nitrogen versions of CLM4.

1.2.4 Nitrogen Model for Assessing the Environmental Impact

The development of N models for environmental and agricultural purposes has involved more than three decades of effort. Based on their model structures and temporal and spatial scales, nutrient models can be categorized into two types: (1) empirical or semi-empirical loading function models and (2) process-based simulation models. Loading function models include the Generalized Watershed Loading Function (GWLF, Haith and Shoemaker, 1987), the Spatially Referenced Regressions on Watersheds (SPARROW, Schwarz et al., 2006), and some regression models; Process-based simulation models include the Agricultural Non-Point Source Pollution Model (AGNPS, Young et al., 1989), the Hydrologic Simulation Program-Fortran (HSPF, Bicknell et al., 1997), and SWAT, all of which are mainly developed in the U.S. (U.S. EPA, 2008). Although loading function models require limited data and are easy to use, they usually fail to replicate the nonlinear biogeochemical and physical processes. In addition, they have limitations in representing spatial and temporal information. In contrast, process-based simulation models incorporate our understanding of physical, chemical, and biological processes, and hence are capable of quantifying the interactions among embedded components (e.g., atmosphere, biosphere, hydrosphere, nutrients, and human activities). However, they tend to require detailed data and have large uncertainties about (1) model structure that is built based on prior

6

knowledge, (2) model parameters, (3) bias from observational data, and (4) users' skills and experiences (U.S. EPA, 2008). Besides these U.S. models, some models have been developed across Europe (Kronvang et al., 2009; Schoumans et al., 2009b), for example, the Nutrient Losses at catchment scale model (NL-CAT, Schoumans et al., 2009a), and the Transport–Retention–Källfördelning N leaching model (TRK, Brandt and Ejhed, 2002). Similar to SWAT, the more complex processed-based models like NL-CAT and TRK are designed to assess the impact of nutrient management strategies on the nutrient losses to the environment under different conditions; however, none of these nutrient models is suitable for all regions of the globe and for all purposes (Schoumans et al., 2009b).

1.3 RESEARCH OBJECTIVES

This dissertation uses the Noah-MP LSM (Niu et al., 2011; Yang et al., 2011) as an example. Noah-MP incorporates recent model development in parameterizations of dynamic leaf, groundwater, and multilayer snow. It is the next generation LSM for the operational weather and climate models in NCEP, and is used from day-to-day weather forecasts to decadal climate projection. Noah-MP is also part of the Weather Research and Forecasting (WRF) meteorological model (Rasmussen et al., 2014), which is used not only for weather forecasts but also for regional applications such as water resources management and climate change impacts. Specifically, the objectives of this dissertation are to: (1) understand the strength and limitation of Noah-MP in characterizing the hydrological cycle; (2) evaluate our confidence about current LSMs in modeling the water balance at the continental scale; (3) develop an integrated modeling capability for simultaneous climate and environmental predictions, and (4) investigate the effects of nitrogen dynamics on carbon and water cycles.

Correspondingly, this dissertation includes three main chapters, followed by a conclusion chapter. Chapter 2 evaluates Noah-MP's performance in simulating the hydrological cycle for the Mississippi River Basin. Model simulated runoff, groundwater, evapotranspiration, soil moisture, and terrestrial water storage are compared against observed data sets. Chapter 3 compares four LSMs' performance in assessing the water balance over the conterminous U.S. (CONUS), based on the North American Land Data Assimilation System (NLDAS) test bed. Chapter 3 is different from Chapter 2 in: (1) it employs the tools (e.g., the river routing model) and data sets (e.g., the streamflow for 961 small river basins) that were developed within the NLDAS test bed; (2) four major LSMs are compared, which gives a better picture showing each model's position in the land model community; and (3) it covers a larger study domain and longer time period. Chapter 4 describes the nitrogen dynamic model and its coupling with Noah-MP. The model simulated nitrogen state/flux variables are evaluated at an experimental site. In addition, the impacts of the nitrogen dynamic module on carbon and water cycles are analyzed. Chapter 5 summarizes the conclusions and findings from this dissertation. The chapter ends with directions for future work.

Chapter 2: Hydrological Evaluation of the Noah-MP Land Surface Model for the Mississippi River Basin[1]

2.1 ABSTRACT

This study evaluates regional-scale hydrological simulations of the newly developed community Noah land surface model (LSM) with multiparameterization options (Noah-MP). The model is configured for the Mississippi River Basin and driven by the North American Land Data Assimilation System Phase 2 atmospheric forcing at 1/8° resolution. The simulations are compared with various observational data sets including the U.S. Geological Survey streamflow and groundwater data, the AmeriFlux tower micrometeorological evapotranspiration (ET) measurements, the Soil Climate Analysis Network (SCAN)-observed soil moisture data, and the Gravity Recovery and Climate Experiment satellite-derived terrestrial water storage (TWS) anomaly data. Compared with these observations and to the baseline Noah LSM simulations, Noah-MP shows significant improvement in hydrological modeling for major hydrological variables (runoff, groundwater, ET, soil moisture, and TWS), which is very likely due to the incorporation of some major improvements into Noah-MP, particularly an unconfined aquifer storage layer for groundwater dynamics and an interactive vegetation canopy for dynamic leaf phenology. Noah-MP produces soil moisture values consistent with the SCAN observations for the top two soil layers (0–10 cm and 10–40 cm), indicating its great potential to be used in studying land–atmosphere coupling. In addition, the simulated groundwater spatial patterns are comparable to observations; however, the inclusion of groundwater in Noah-MP requires a longer spin-up time, ranging from less than 10 years for wet regions to hundreds of years for arid regions. Runoff simulation is highly sensitive

---

to three parameters: the surface dryness factor (α), the saturated hydraulic conductivity (k), and the saturated soil moisture (θ_{max}). This study identifies groundwater–river interaction, leaf dynamics, and agricultural modeling as areas for further research.

2.2 INTRODUCTION

Land surface models (LSMs) have evolved rapidly in recent decades due to the advances in high-performance computing, ground-based measurements (e.g., FLUXNET (Baldocchi et al., 2001)), remote sensing (Murray et al., 2013), and emerging concepts such as hyperresolution (Wood et al., 2011) and multiparameterization (or multiple hypotheses) (Clark et al., 2011b). One such LSM is the community Noah LSM with multiparameterization options (hereafter Noah-MP) (Niu et al., 2011; Yang et al., 2011). Based on the Noah LSM (Ek et al., 2003), Noah-MP has added biophysical processes such as an unconfined aquifer for groundwater storage and a dynamic water table (Niu et al., 2007), an interactive vegetation canopy (Dickinson et al., 1998), a multilayer snowpack (Yang and Niu, 2003), and a simple TOPMODEL (TOPography based hydrological MODEL)-based runoff production (Niu et al., 2005).

Model evaluation plays a very important role in LSM development, as LSM benchmarking or better model evaluation is one of the three core activities in the current Global Energy and Water Cycle Exchanges Project (GEWEX) Global Land/Atmosphere System Study (GLASS) (van den Hurk et al., 2011). Noah-MP has been tested at local scales (Niu et al., 2011) and in global river basins (Yang et al., 2011). Its runoff simulation was evaluated using the University of New Hampshire-Global Runoff Data Center (UNH-GRDC) gridded runoff data set (Fekete et al., 2002), but it has not yet been evaluated with the U.S. Geological Survey (USGS) streamflow data. The groundwater module was evaluated against the Gravity Recovery and Climate Experiment (GRACE) terrestrial

water storage (TWS) data when it was coupled with the Community Land Model (Niu et al., 2007), but it has not been evaluated since its coupling with Noah-MP. In addition, the evapotranspiration (ET) simulation has not yet been evaluated using observational data. Soil moisture is another important variable for land–atmosphere coupling and drought monitoring and thus needs to be evaluated using observational data. Lastly, since the launch of GRACE in 2002, modeled TWS can now be evaluated at a regional scale with the GRACE-derived TWS.

Therefore, this study comprehensively evaluates the performance of Noah-MP in hydrological simulations of major hydrological variables (runoff, groundwater, ET, soil moisture, and TWS). It features a detailed multivariable evaluation using the best available ground-based and satellite measurements. This type of evaluation is consistent with the current call for benchmarking of LSMs by GEWEX GLASS (van den Hurk et al., 2011; Abramowitz, 2012; Kumar et al., 2012). Specifically, the following measurements are used: USGS streamflow and groundwater data, AmeriFlux tower micrometeorological ET data, Soil Climate Analysis Network (SCAN) soil moisture data, and GRACE satellite-derived TWS anomaly data.

Before a clean version of the model is obtained for evaluation, hydrological modeling generally requires spin-up, parameter sensitivity tests, and model calibration for specific study areas. Without groundwater dynamics, hydrological models typically only require several years of model spin-up time (Cosgrove et al., 2003b); with groundwater dynamics, however, they require much longer time for water table depth (WTD) to reach an equilibrium state, particularly in arid regions (Niu et al., 2007). This study will first investigate how the integration of groundwater dynamics into Noah-MP affects model spin-up. Unlike some parameters (e.g., slope and leaf area index) that can be derived from representative field site measurements or remote-sensing data, many parameters, such as

11

the hydraulic conductivity and the Clapp-Hornberger "b" parameter (a fitting parameter in determining soil water potential from hydraulic conductivity), cannot be directly derived from measurements and must be estimated by calibration for the specific study areas (Hogue et al., 2006). This study will identify the most sensitive parameters through sensitivity tests and then obtain the optimal combination of these parameters through calibration.

The evaluation of the study is conducted for the period of 2000–2009 for the Mississippi River Basin (MRB) at the North American Land Data Assimilation System Phase 2 (NLDAS-2)'s 1/8° resolution (Ek et al., 2011). Section 2 describes the Noah-MP model, study area, and data sets used in this study. Section 3 introduces model spin-up, parameter sensitivity tests, and model calibration. Section 4 shows the specific evaluations of runoff, groundwater, ET, soil moisture, and TWS. Section 5 summarizes the study.

2.3 METHODOLOGY

2.3.1 The Noah-MP Model

Noah-MP was enhanced from the original Noah LSM through the addition of improved physics and multiparameterization options (Niu et al., 2011; Yang et al., 2011). The improved physics includes a dynamic groundwater component, an interactive vegetation canopy, and a multi-layer snowpack. The multiparameterization options provide users with multiple choices of parameterizations in leaf dynamics, canopy stomatal resistance, soil moisture factor for stomatal resistance, and runoff and groundwater. For example, there are four options for runoff and groundwater: (1) TOPMODEL with simple groundwater model (SIMGM) (Niu et al., 2007), (2) TOPMODEL with an equilibrium water table (Niu et al., 2005), (3) original surface and subsurface runoff (free drainage) (Schaake et al., 1996), and (4) Biosphere-Atmosphere Transfer Scheme surface and

subsurface runoff (free drainage) (Yang and Dickinson, 1996). The parameterizations used in this study are the default options recommended by Yang et al. (2011): TOPMODEL runoff with SIMGM groundwater, leaf dynamics, Ball-Berry canopy stomatal resistance, a Noah-type (using soil moisture) soil moisture factor controlling stomatal resistance, and the Monin-Obukhov surface exchange coefficient for heat.

Both surface and subsurface runoff are computed by a simple TOPMODEL-based runoff model (Niu et al., 2005). Surface runoff (R_{sf}) is parameterized as

$$R_{sf} = F_{sat}\, p + (1 - F_{sat})\max(0, (p - I))$$

(2.1)

where p, the effective precipitation intensity (kg m^{-2} s^{-1}), is the rainfall and dewfall reaching the ground plus snowmelt, I is maximum soil infiltration capacity (kg m^{-2} s^{-1}), which is dependent on soil properties and moisture, and F_{sat} is the fraction of saturated area and is parameterized as

$$F_{sat} = (1 - F_{frz})F_{max}\, e^{-0.5 f z_\nabla} + F_{frz}$$

(2.2)

where F_{frz} is a fractional impermeable area as a function of the soil ice content of the surface soil layer, z_∇ is the WTD (m), and F_{max} is the potential or maximum saturated fraction for a grid cell, which can be derived from high-resolution subgrid topography (e.g., 30 m) of a model grid cell (e.g., 1° resolution) using the TOPMODEL concepts (see Niu et al. (2005) or Niu et al. (2011) for details). In this study, a global mean $F_{max} = 0.38$ is used, which is derived from the HYDRO1K 1 km topographic index (or wetness index, WI) data (Verdin and Jenson, 1996).

Subsurface runoff (R_{sb}) is parameterized as

$$R_{sb} = R_{sb,max}\, e^{-\Lambda - f z_\nabla}$$

(2.3)

13

where $R_{sb,max}$ is the maximum subsurface runoff when the grid cell mean WTD is zero—here globally $R_{sb,max} = 5.0 \times 10^{-4}$ m s^{-1}, derived from calibration against global runoff data through sensitivity tests (Niu et al., 2007)—and Λ is the grid cell mean WI—here $\Lambda =$ 10.46, which is its global mean value derived from HYDRO1K 1 km WI data.

With an unconfined aquifer added to account for the exchange of water between the soil and the aquifer, the temporal variation of the water stored in the unconfined aquifer, W_a (mm), is parameterized as

$$\frac{\mathrm{d}W_a}{\mathrm{d}t} = Q - R_{sb} \tag{2.4}$$

where Q is the recharge rate (mm s^{-1}), which is positive when water enters the aquifer. It is parameterized as

$$Q = -K_{bot} \frac{-z_\nabla - (f_{mic}\psi_{bot} - z_{bot})}{z_\nabla - z_{bot}} \tag{2.5}$$

where K_{bot} is hydraulic conductivity of the bottom soil layer (mm s^{-1}). f_{mic} is the fraction of micropore content in the bottom-layer soil, which is introduced to limit the upward flow (depending on the level of structural soil) and ranges from 0.0 to 1.0 (see Niu et al. (2011) for details); and ψ_{bot} is the matric potential (mm).

Latent heat flux (λE, or potential evapotranspiration E) is calculated using the Penman-Monteith equation following Bonan (2008):

$$\lambda E = \frac{s(R_n - G) + \rho C_p (e_*[T_a] - e_a)/r_H}{s + \gamma(r_W / r_H)} \tag{2.6}$$

where λ is the latent heat of vaporization (J kg^{-1}), $e_*[T_a]$ is the saturation vapor pressure evaluated at the air temperature (T_a); $s = \mathrm{d}e_*[T_a]/\mathrm{d}T$ is the saturation vapor pressure versus temperature evaluated at T_a, R_n is net radiation (W m^{-2}), G is soil heat flux (W m^{-2}), (R_n –

14

G) is net available radiation (W m^{-2}), ρ is dry air density (kg m-3), C_p is specific heat capacity of air (J kg^{-1} K^{-1}), e_a is the vapor pressure of air (Pa), and r_H and r_W are resistance to sensible heat and water vapor, respectively (s m^{-1}). The surface exchange coefficient for heat, C_H, which is used to calculate aerodynamic resistances, can be estimated using either the Monin–Obukhov similarity theory (this study) or the method of Chen et al. (1997).

In addition to hydraulic conductivity, runoff is also found to be very sensitive to the surface dryness factor (α). It determines the soil surface resistance to ground evapotranspiration (Sellers et al., 1992), as shown in the following equation:

$$r_{surf} = f_{snow} \times 1.0 + \left(1 - f_{snow}\right) e^{(8.206 - \alpha S_1)} \tag{2.7}$$

where r_{surf} is the soil surface resistance (s m^{-1}), f_{snow} is the snow fraction covering a ground surface, and S_1 is the soil wetness in the top soil layer, varying from 0.0 to 1.0. Thus, α controls the effect of soil moisture on r_{surf}.

In the current Noah-MP, vegetation plays a significant role in the model: the stomatal conductance determines the photosynthesis and the carbon cycle; the dynamic leaf model predicts the leaf area index (LAI) and the green vegetation fraction (GVF); the "semitile" subgrid scheme calculates the surface energy balance for vegetated and bare ground separately; and the canopy water scheme simulates the canopy water interception and evaporation.

2.3.2 Study Area

The Mississippi River Basin (MRB) is the largest river basin in North America, covering many distinct climate zones. It is also a well-studied river basin, and thus, a variety of meteorological, hydrological and ecological data are available. For example, it is the study area of the first Continental-Scale Experiment of the World Climate Research Program GEWEX Continental-Scale International Project (Roads et al., 2003; Kumar and

Merwade, 2011). The MRB area is 3.28 million km^2, which is approximately 41% of the conterminous U.S. (Figure 2.1a). It covers six of the 21 major geographic regions defined by the USGS two-digit hydrologic unit code (HUC2, http://water.usgs.gov/GIS/huc.html). Calculated from the NLDAS-2 meteorological forcing data (1998–2009), the basin average annual temperature and precipitation are 11.9°C and 821.0 mm, respectively. Across the various climate zones, there is a large temperature gradient between the south and north, with a minimum of −3°C in the Rocky Mountains and a maximum of 22.9°C in the southern most area of the basin (Figure 2.1b), and a large precipitation gradient between the southeast and northwest, with a minimum of 126.3 mm in Wyoming and a maximum of 1973.6 mm in the Gulf of Mexico region (Figure 2.1c). In this study, the Ohio-Tennessee Region is considered a typical wet region and Missouri Region a typical dry region, with the Upper Mississippi Region considered the transitional zone between the two.

2.3.3 Model Input Data

The NLDAS-2 (Mitchell et al., 2004) meteorological forcings at 0.125° spatial resolution and hourly temporal resolution are used to drive the Noah-MP model. The seven non-precipitation meteorological forcing fields are derived from the NCEP (National Centers for Environmental Prediction) North American Regional Reanalysis, including air temperature, the U and V components of wind speed, specific humidity, surface pressure, surface downward shortwave radiation, and surface downward longwave radiation. Precipitation field data are derived from the temporal disaggregation of the gaged daily precipitation data from NCEP/Climate Prediction Center with an orographic adjustment based on the monthly climatological precipitation of the Parameter-elevation Regressions on Independent Slopes Model (Daly et al., 1994). The Noah LSM outputs forced by the same NLDAS-2 meteorological forcings are also downloaded from the NLDAS website,

16

which serves as the baseline model for comparison with Noah-MP. More details regarding the setup and performance of the Noah LSM model can be found in previous studies (Xia et al., 2012a; Xia et al., 2012c).
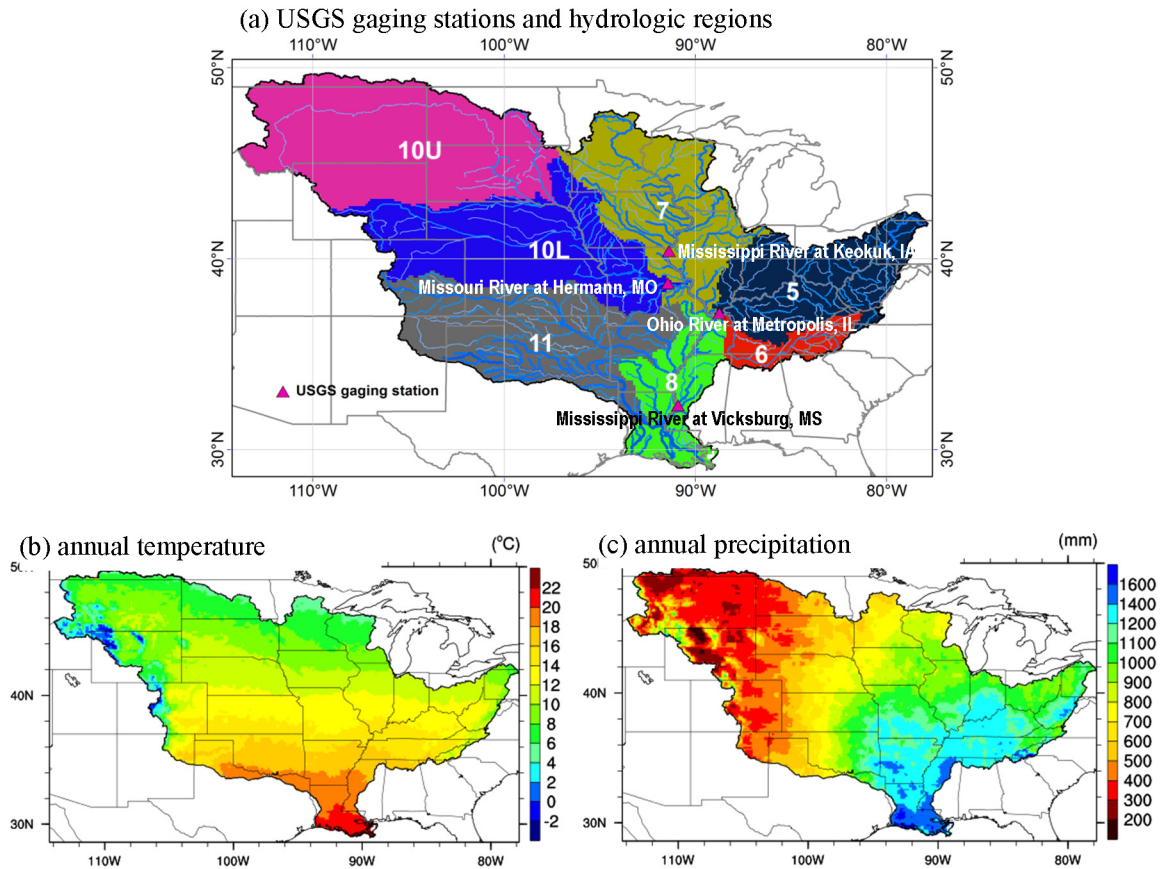
Figure 2.1 Map of the Mississippi River Basin showing (a) USGS gaging stations and hydrologic regions (Numbers in the shaded area are the 2-digit hydrologic unit code: 5–Ohio; 6–Tennessee; 7–Upper Mississippi; 8–Lower Mississippi; 10U–Upper Missouri; 10L–Lower Missouri; 11–Arkansas–White–Red); (b) Annual temperature; and (c) Annual precipitation.

The static input data for Noah-MP are from various sources. The land-water mask that masks out the water component from simulation (land = 1 and water = 0), and the latitude and longitude coordinate information, which are primarily used for computing the

solar zenith angle, are the same as those of NLDAS-2, which uses the latitude and longitude in the center of each 0.125° grid box. The vegetation type and soil texture types (top 30 cm and 30–100 cm depth) are aggregated from the 30 arc-second data of the USGS 24-category vegetation (land use) and the hybrid State Soil Geographic Database (STATSGO) Food and Agriculture Organization soil texture data sets, respectively, both of which are maintained by the NCAR/RAL (Research Application Laboratory, National Center for Atmospheric Research) (http://www.ral.ucar.edu/research/land/technology/lsm.php). Soil color data are used to determine ground surface albedo over visible and infrared bands and include eight categories, with one as the lightest and eight as the darkest. The annual mean 2 m air temperature data (from NCAR/RAL) are also used as the lower boundary layer condition for soil temperature. The monthly climatological GVF data are converted from the 0.144° five year (1985–1990) GVF data derived from National Oceanic and Atmospheric Administration (NOAA)/advanced very high resolution diameter (AVHRR) by Gutman and Ignatov (1998).

2.3.4 Observational Data

USGS streamflow data are used for runoff calibration and validation. As shown in Figure 2.1a, four USGS gaging stations are selected: the Ohio River at Metropolis, IL collecting runoff from region 5 and region 6; the Mississippi River at Keokuk, IA collecting runoff from region 7; the Missouri River at Hermann, MO collecting runoff from region 10; and the Mississippi River at Vicksburg, MS collecting runoff from the entire MRB. To compare the spatial distribution of runoff between the model and the observations, we use two types of data sets: the USGS hydrologic unit runoff (http://waterwatch.usgs.gov/) (Brakebill et al., 2011) and the monthly gridded climatological runoff composite fields at 30 min spatial resolution provided by the University of New Hampshire-Global Runoff

18

Data Center (UNH-GRDC). The USGS hydrologic unit runoff data set was developed in 2008 and has been updated annually (D. Wolock, personal communication, 2012). It was calculated from all the available records for 1901–2009 at the eight-digit hydrologic unit code (HUC8) level, which consists of 2110 hydrologic cataloging units for the continental U.S. and 1128 units for the MRB. The UNH-GRDC runoff preserves the accuracy of the observed discharge data and obtains consistent spatial and temporal resolutions from a water balance model (Fekete et al., 2002); hence, it is considered the best available gridded data set for LSM evaluation, although the values are occasionally lower than the gaged discharge data (Leung et al., 2011).

A climatological WTD map for the U.S. created by Fan and Miguez-Macho (2011), which contains 567,946 USGS groundwater observational sites (254,464 sites for the MRB) and is dense enough to show the groundwater spatial pattern for most of the U.S., is used for spatial comparison with the Noah-MP simulated WTD. Daily groundwater storage anomalies for the MRB and its major subbasins were derived from 58 sites with good representation of the subbasin averages by Rodell et al. (2007) and are used in this study for temporal comparison with the model results.

Over recent decades, the global network of micrometeorological tower sites with coordinating eddy covariance measurements of CO_2, water vapor, and energy, the FLUXNET (http://fluxnet.ornl.gov/) (Running et al., 1999; Baldocchi et al., 2001), has provided the most reliable ET measurements and has been considered a valuable data source for LSM development (Stöckli et al., 2008) and evaluation (Blyth et al., 2010; Kumar and Merwade, 2011; Li et al., 2011). As part of the FLUXNET, AmeriFlux features much denser flux tower sites in the U.S. than other regional networks. Although there are 21 AmeriFlux tower sites in the MRB, only 15 of them have overlapping observation times with the Noah-MP simulation period (2000–2009), and those sites are used to evaluate the

model-simulated latent heat flux (ET). A list of Ameriflux tower sites and their locations, land cover, and available measurement periods are shown in Table 2.1. The data included in this study are the monthly Level 4 latent heat flux data, which are gap filled using Artificial Neural Network and Marginal Distribution Sampling techniques (http://public.ornl.gov/ameriflux).

The SCAN (Schaefer et al., 2007) is a nationwide soil moisture and climate information system led by the Natural Resources Conservation Service (NRCS), USDA (http://www.wcc.nrcs.usda.gov/scan/). SCAN soil moisture data are collected by dielectric constant measuring devices at 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm, where possible. The data used in this study have been processed by extensive quality control steps (Liu et al., 2011), through which any unrealistic data values (e.g., data outside a reasonable range and inconsistent data affected by sensor calibration or installation) and data measured under frozen conditions were excluded. Figure 2.2 shows the 60 available SCAN stations in the MRB and their data availability. Due to the low data availability at most stations, the data are aggregated into monthly basin average.

Table 2.1 Basic information regarding the AmeriFlux sites.

Site no.	Site name	State	Latitude (°N)	Longitude (°E)	Land cover[a]	Available period
1	ARM Southern Great Plains burn site-Lamont	OK	35.55	−98.04	GRA	2005–2006
2	ARM Southern Great Plains control site- Lamont	OK	35.54	−98.04	GRA	2005–2006
3	ARM Southern Great Plains site-Lamont	OK	36.61	−97.49	CRO (DCP)	2003–2006
4	Brookings	SD	44.35	−96.84	GRA	2004–2006
5	Bondville	IL	40.01	−88.29	CRO (DCP)	1996–2007
6	Fort Peck	MT	48.31	−105.10	GRA	2000–2006
7	Goodwin Creek	MS	34.25	−89.97	GRA	2002–2006
8	Lost Creek	WI	46.08	−89.98	CSH (SHR)	2001–2005
9	Morgan Monroe State Forest	IN	39.32	−86.41	DBF	1999–2006
10	Missouri Ozark Site	MO	38.74	−92.20	DBF	2004–2007
11	Mead – irrigated continuous maize site	NE	41.17	−96.48	CRO (DCP)	2001–2006
12	Mead – irrigated maize-soybean rotation site	NE	41.16	−96.47	CRO (DCP)	2001–2006
13	Mead – rainfed maize-soybean rotation site	NE	41.18	−96.44	CRO (DCP)	2001–2006
14	Niwot Ridge Forest (LTER NWT1)	CO	40.03	−105.55	ENF	1998–2007
15	Willow Creek	WI	45.81	−90.08	DBF	1999–2006

[a] AmeriFlux uses IGBP land cover classification, while Noah-MP uses USGS global 24-category classification. GRA stands for Grassland; CRO: Cropland, DCP: Mixed Dryland/Irrigated Cropland and Pasture; CSH: Closed Shrublands; SHR: Shrubland; DBF: Deciduous Broadleaf Forest; ENF: Evergreen Needleleaf Forest. Unless otherwise indicated by parentheses, Noah-MP uses the same land cover type as AmeriFlux for the corresponding 0.125°×0.125° model grid. There are six sites where Noah-MP and AmeriFlux use slightly different names. At site 3, for example, AmeriFlux uses CRO, while Noah-MP uses DCP.
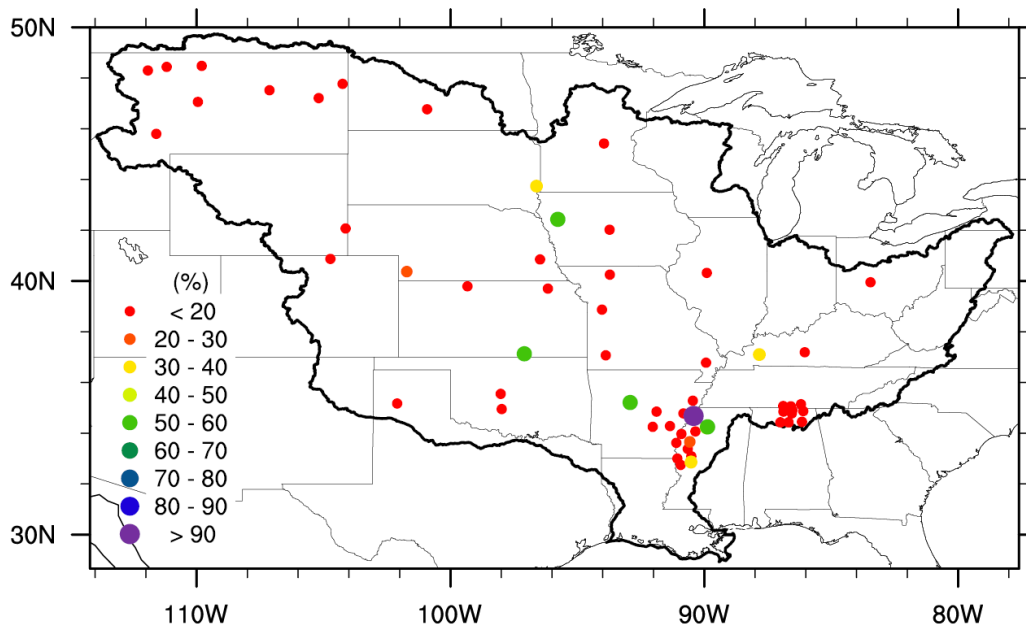
Figure 2.2 Map of the 60 SCAN stations and their data availability (percentage of the total number of months from 2002 to 2007 with observational data) in the Mississippi River Basin.

Section 1.2.2 has described the GRACE TWS data. This study uses the monthly GRACE TWS anomaly data, which have been processed into a 1°×1° resolution gridded format (Swenson and Wahr, 2006; Landerer and Swenson, 2012) for easy comparison with LSM outputs and which can be publicly accessed on the Jet Propulsion Laboratory TELLUS website (http://grace.jpl.nasa.gov). The data are based on the CSR RL4.0 release by the Center for Space Research at the University of Texas at Austin. First, a destriping filter was applied to the data to minimize the systematic errors, which manifest as north-south oriented "stripes" in the GRACE TWS maps; then a 300 km wide Gaussian filter was applied to reduce random errors in higher-degree spherical harmonic coefficients not removed by the previous filter; and lastly, a spherical harmonic filter cutoff at 60° was applied. During the filtering process, because GRACE TWS was spatially averaged, signals were attenuated by showing smaller root-mean-square variability. To restore the

signal attenuation, a gain factor, which was derived by using a simple least square regression to minimize the mismatch between the unfiltered, true, and filtered storage time series, was applied to each of the 1°×1° grids. More information about the data processing can be found in Landerer and Swenson (2012), Chen et al. (2006), and Swenson and Wahr (2006).

2.3.5 Evaluation Statistics

The agreement between the values predicted by a model and the values actually observed is measured using the following statistics: mean, root-mean-square error (RMSE), square of the correlation coefficient (R^2), and Nash–Sutcliffe efficiency (NSE) coefficient (Nash and Sutcliffe, 1970). The NSE is calculated as

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}(M_i - O_i)^2}{\sum_{i=1}^{N}(O_i - \overline{O})^2} \tag{2.8}$$

where M_i and O_i are the predicted and measured values of the same variable, respectively, and $\overline{O}$ is the mean of the measured values. NSE ranges from minus infinity (poor fit) to 1 (perfect fit). In general, model prediction is considered to be satisfactory if NSE > 0.50 (Moriasi et al., 2007).

2.4 MODEL SPIN-UP, SENSITIVITY TESTS, AND CALIBRATION

2.4.1 Model Spin-Up

To allow some of the model variables with longer memories reach equilibrium, a numerical model must be properly initialized. When SIMGM, the groundwater component of Noah-MP, was introduced (Niu et al., 2007), it was noted that it might take at least 250 years to spin-up the WTD in arid regions. Therefore, we are interested in examining the

23

time span required to spin-up the model for this river basin. In this study, the spin-up is completed by running the model repeatedly through 1997 until each of the variables reaches equilibrium and the spin-up time is defined as year *n*, if

$$\left| Var^{\,n+1} - Var^{\,n} \right| < 0.001 \cdot \left| Var^{\,n} \right|$$

$$(2.9)$$

where *Var* stands for each of the variables for the spin-up. This definition is as strict as the constraint by Yang et al. (1995). The *Var* for the calculation in Figure 2.3a is spatially averaged, and for the calculation in Figure 2.3b, it is averaged per grid cell.

The WTD requires the longest spin-up time, 34 years (Figure 2.3a), followed by runoff with 11 years and soil moisture (total soil column) with 8 years. This is consistent with previous studies of WTD (Niu et al., 2007) and soil moisture (Yang et al., 1995; Cosgrove et al., 2003b). However, it is surprising to note that the spin-up time needed for runoff is longer than for soil moisture, which is not as commonly reported in literature. This may be due to the long spin-up time for WTD, which influences the runoff generation. The sensible heat fluxes and latent heat fluxes need shorter times to spin-up, approximately 4 years, because they are more influenced by surface soil and vegetation states and by the atmospheric forcing data, as indicated in Equation 2.6.

Regarding the spatial distribution of the time (in years) required for WTD to reach equilibrium (Figure 2.3b), for the wet region (east), less than 10 years is required to spin-up, whereas for the dry region (west), more than 72 years or even hundreds of years may be required for some small but extremely dry areas.

24
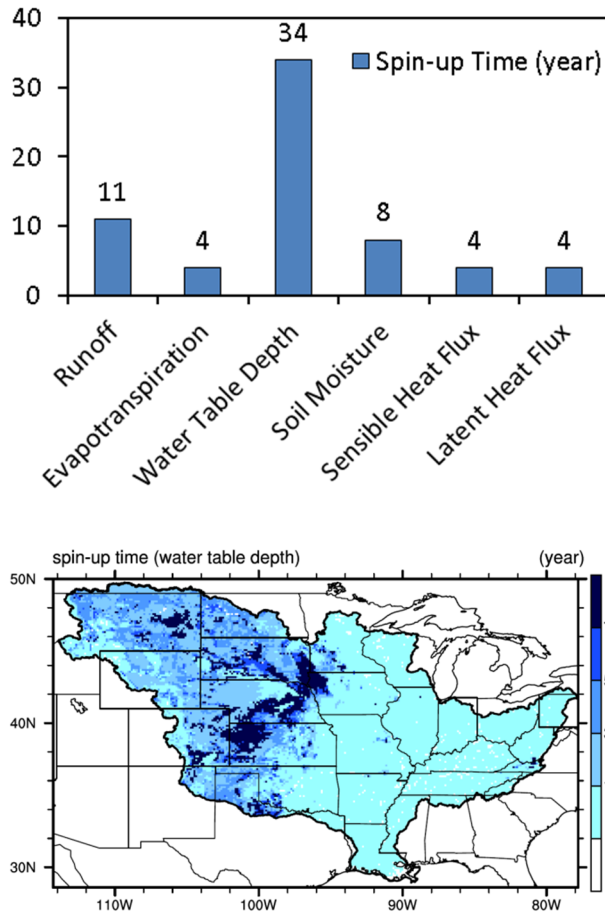
Figure 2.3 (a) Spin-up time (in years) for the individual variables based on averaged values for the entire Mississippi River Basin; (b) Spatial distribution of the spin-up time (in years) for the water table depth.

2.4.2 Parameter Sensitivity Tests

Hydrological modeling involves significant efforts in parameter sensitivity testing and calibration, which were usually overlooked in the past. However, it is becoming a must when LSMs are more and more used in hydrological studies. Here we briefly describe how the model parameters are finalized before the model is ready for evaluation. Based on our modeling experience and previous studies (Rosero et al., 2010), several sensitive parameters are selected for further analysis. However, only three parameters are identified as sensitive parameters for runoff simulation: surface dryness factor (α), saturated

hydraulic conductivity (k), and saturated soil moisture (θ_{max}). Table 2.2 provides the definitions, units, and value ranges of the three parameters. Figure 2.4 shows how annual runoff varies with different values of the individual parameters. Spatially-averaged annual mean runoff (1) decreases as the surface dryness factor increases, in a nearly linear relationship; (2) decreases dramatically as hydraulic conductivity increases when saturated hydraulic conductivity is less than 10% of its original values; and (3) also decreases dramatically as saturated soil moisture increases when the multiplier (a factor multiplied by the original values) is less than 0.9.

Although the sensitivity tests are based on the changes in total runoff with different parameter values, they influence other hydrological variables more directly. For example, as indicated in Equation 2.7, the surface dryness factor is a parameter that determines the soil surface resistance and hence controls soil evaporation; when the surface dryness factor increases, the soil evaporation increases correspondingly. Therefore, to maintain water balance, either or both transpiration and runoff have to decrease. In this case, annual runoff decreases as the surface dryness factor increases (Figure 2.4a). Total runoff is affected by the saturated hydraulic conductivity via its capability to control the subsurface runoff (base flow). Similarly, saturated soil moisture controls the storage capacity of the soil and hence affects evaporation and infiltration.

Table 2.2 Experimental design for parameter calibration.

Parameters	Controlling process	Units	Min	Max	Default	Values	#
Surface dryness factor	partitions of the surface hydrology	—	0	10	6.0	2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6	8
Saturated hydraulic conductivity	base flow in runoff simulation	m s^{-1}	2E-9	7E-2	Vary	Multiply by 0.01, 0.05, 0.1, 0.5, 1, 5, 10	7
Saturation soil moisture content	water flow between aquifer and soil	—	0.10	0.71	vary	Multiply by 0.8, 0.85, 0.9, 0.95, 1, 1.05, and 1.1	7
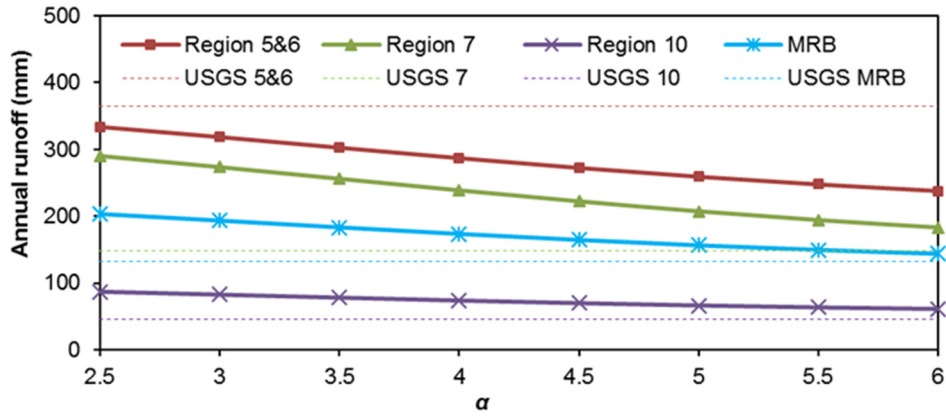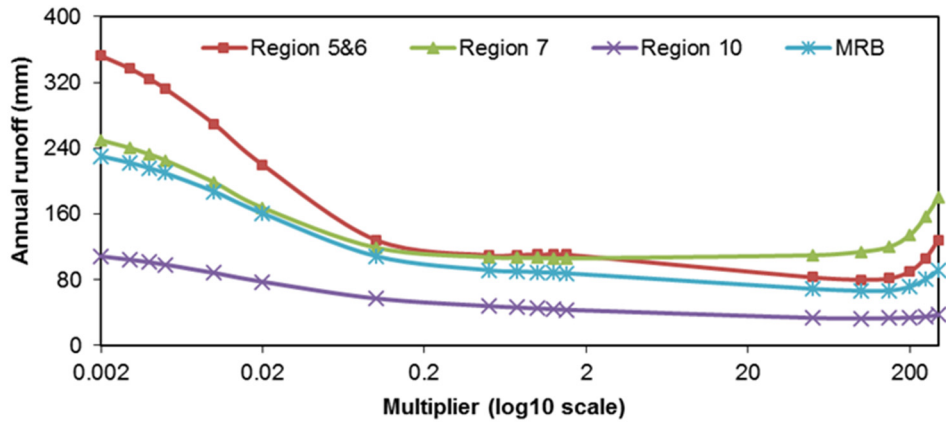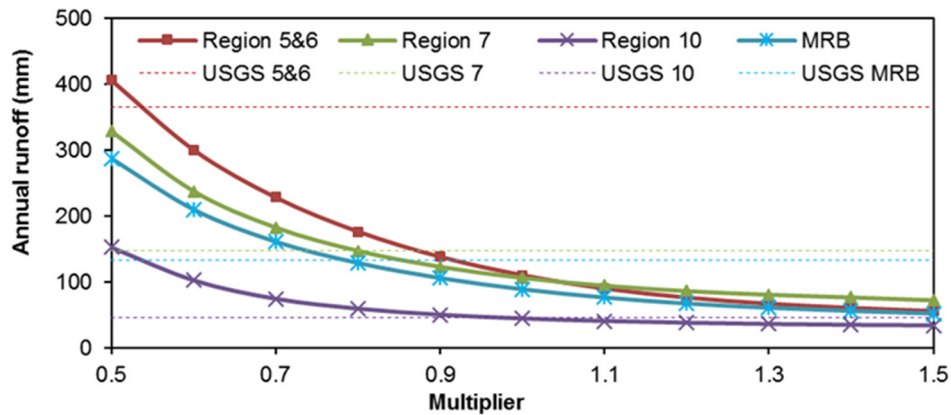
Figure 2.4 Sensitivity tests of (a) the surface dryness factor, (b) the saturated hydraulic conductivity and, (c) the maximum soil moisture (porosity).

2.4.3 Model Calibration

Based on the parameter sensitivity tests above, model calibration is conducted manually by obtaining the optimal combination of the three most sensitive parameters (α, k, and θ_{max}) for the entire MRB (i.e., called lumped calibration). As shown in Table 2.2, 392 experiments are designed and run. The three parameter values that produce the highest NSE for the entire MRB are selected. The calibrated hydrographs are shown in Figure 2.5 and the corresponding statistics are included in Table 2.3. In the hydrograph, the USGS-observed streamflow is from the station near the basin outlet and the Noah-MP runoff is aggregated from all the grids in each basin or subbasins. For the entire MRB, we can observe that the hydrograph is greatly improved from the default simulation to the calibrated simulation. Although the increase in R^2 is small (from 0.76 to 0.81), the decrease in RMSE is large (37%), which results in a large increase in NSE from 0.42 (less than the satisfactory threshold of 0.50) to 0.77. For the Ohio-Tennessee River Basin, all three statistics improve. For the Upper MRB and the Missouri River Basin, however, the NSE decreases (from 0.56 and 0.01 to 0.29 and –0.16, respectively) due to the increase in RMSE, although the R^2 increases.

To improve the simulation for the subbasins, calibration is conducted by using different sets of the three parameters (α, k, and θ_{max}) for each subbasin (i.e., subbasin calibration). However, the improvement to the runoff simulation for the Upper MRB and the Missouri River Basin is very trivial (not shown here). The possible reasons are (1) the selected sensitive parameters are not applicable for these two subbasins, and/or (2) hydrological modeling for arid and semiarid areas (such as these two subbasins) is more difficult than that for humid areas—a well-recognized problem. Figure 2.5d shows that subbasin calibration does significantly improve the simulation for the Ohio–Tennessee River Basin. From the lumped calibration to the subbasin calibration, the NSE (R^2)

28

increases from 0.36 (0.67) to 0.68 (0.81). In summary, it is worthwhile to calibrate models at the subbasin level for humid regions, whereas for arid and semiarid regions, the model structure and the selection of the sensitive parameters need further investigation.

Table 2.3 Statistical summary of model calibration for the Mississippi River Basin and some of its subbasins.

	Mississippi		Upper Mississippi		Missouri		Ohio-Tennessee		
	CTL[a]	LPC[b]	CTL	LPC	CTL	LPC	CTL	LPC	SBC[c]
RMSE	6128	3875	904	1147	1192	1290	5054	3919	2773
R^2	0.76	0.81	0.60	0.68	0.48	0.57	0.60	0.67	0.81
NSE	0.42	0.77	0.56	0.29	0.01	−0.16	−0.07	0.36	0.68

[a] CRT: default (control) model run.
[b] LPC: lumped calibration.
[c] SBC: calibration for specific subbasin.
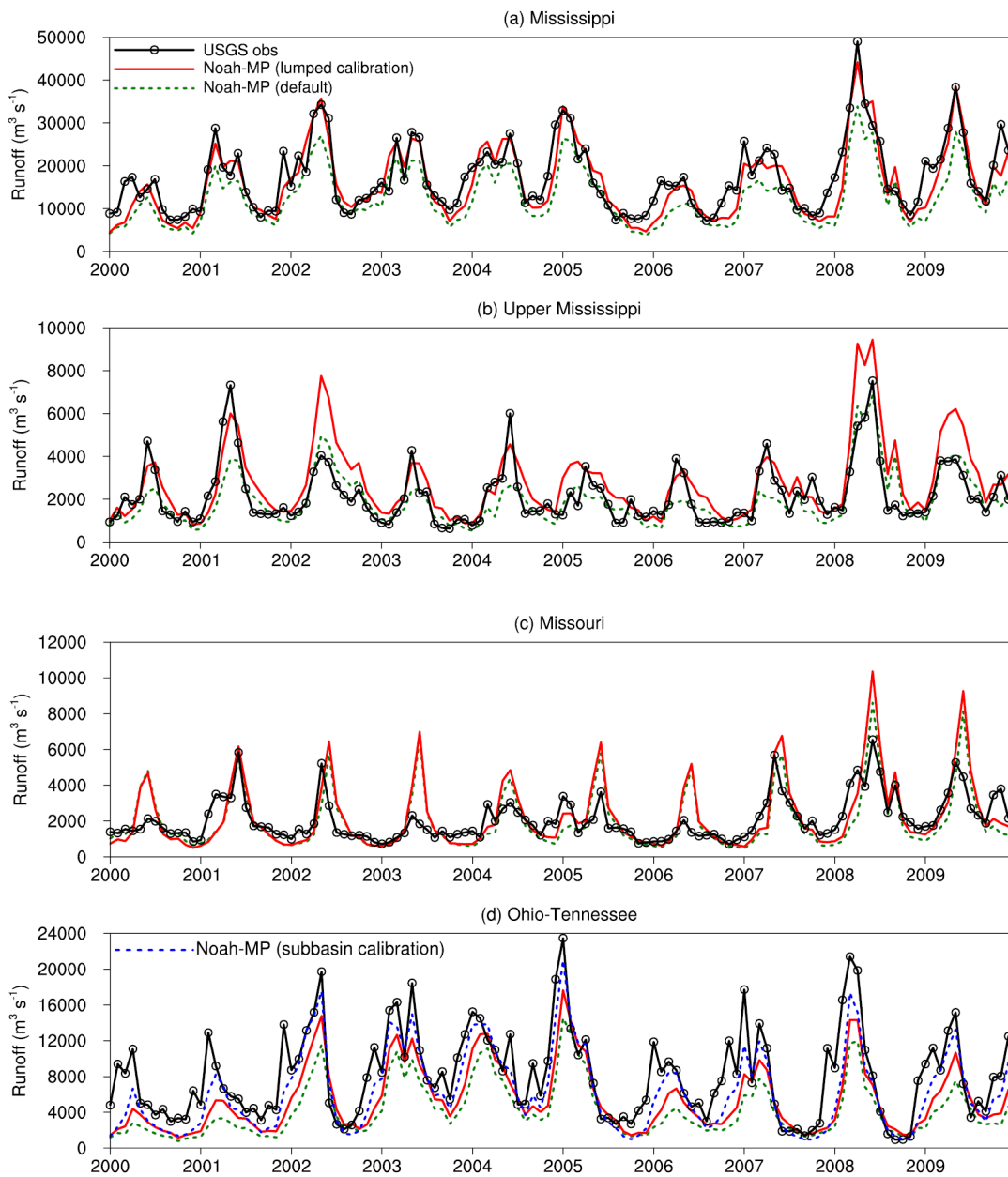
Figure 2.5 Comparisons of the USGS-observed and the Noah-MP- simulated (default, lumped calibrated, and subbasin calibrated) hydrographs for (a) the Mississippi River Basin, (b) the Upper Mississippi River Basin, (c) the Missouri River Basin, and (d) the Ohio–Tennessee River Basin.

2.5 EVALUATION AND DISCUSSION

2.5.1 Runoff

Figure 2.6 compares the spatial distributions of the UNH-GRDC composite runoff, the USGS hydrologic unit runoff, and the Noah-MP simulated runoff. Noah-MP is capable of capturing the observed general spatial pattern of the runoff, which is similar to the precipitation pattern shown in Figure 2.1c. For UNH-GRDC, the runoff in the red box is much lower than its surrounding area, which is not found in the Noah-MP simulation or the USGS observations. On the contrary, the Noah-MP simulated runoff in that box is slightly higher than its surrounding area, whereas the USGS runoff follows the general transition pattern. The USGS stream gages in this region are very sparse compared with other regions (not shown here), which might explain the difference in the UNH-GRDC runoff data because the UNH-GRDC runoff requires runoff input from the USGS. The high Noah-MP runoff in that box corresponds to the high precipitation in Figure 2.1c. It is also notable that the UNH-GRDC runoff is less than 1 mm for most of the western portion of the basin (red shaded area in Figure 2.6b), whereas it is 5 mm to 25 mm in the Noah-MP runoff and 6 mm to 50 mm in the USGS runoff.

31
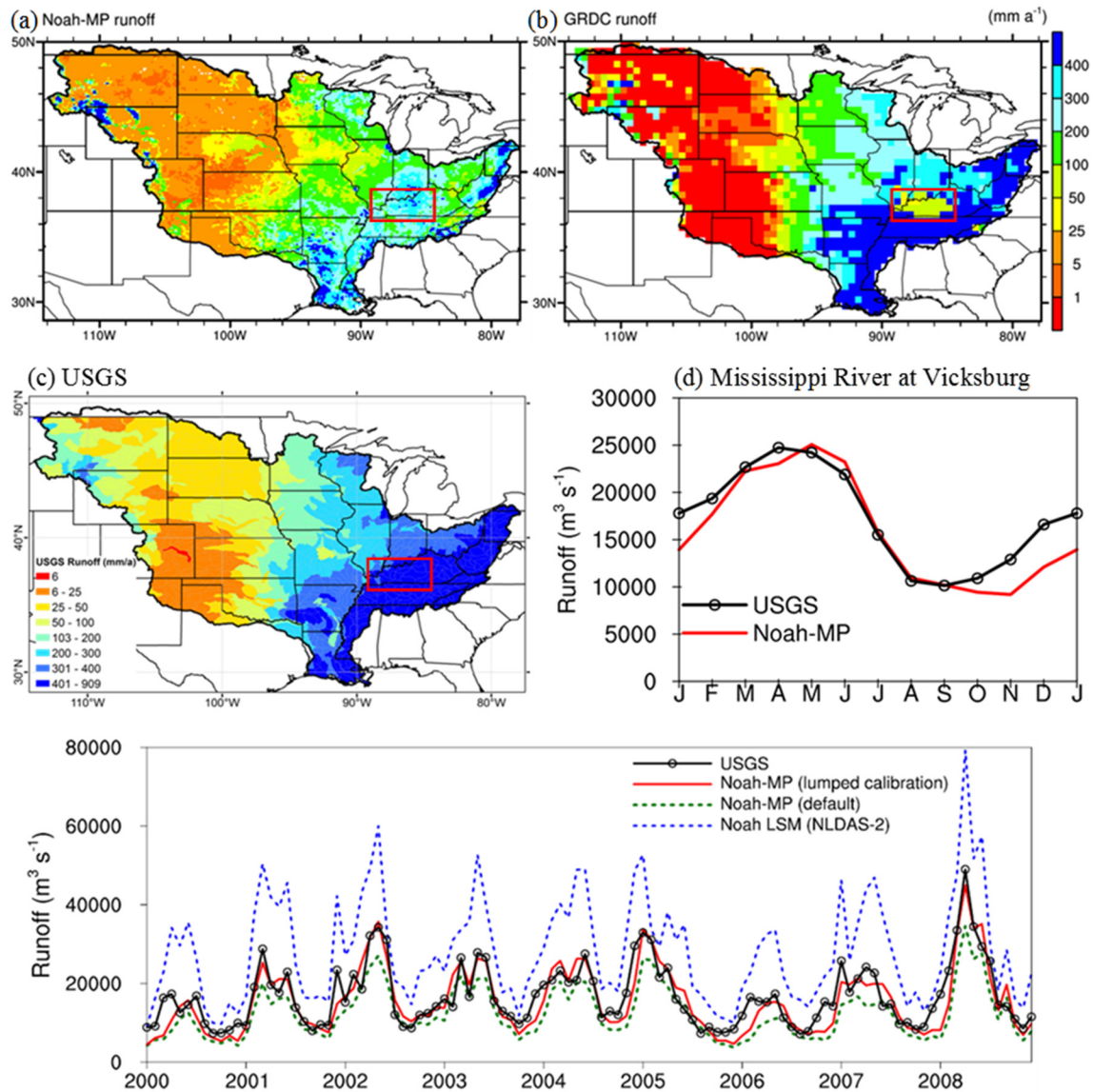
Figure 2.6 Climatological mean annual runoff from (a) Noah-MP (2000–2009), (b) UNH-GRDC (all available observational time periods (Fekete et al., 2002)), and (c) USGS hydrologic unit runoff (1901–2009). (d) Monthly climatological runoff (2000–2009) from USGS observation and Noah-MP simulation. (e) USGS-gauged and Noah-MP (default and calibrated) and Noah LSM simulated runoff (2000–2008) for the Mississippi River Basin. The region enclosed in the red box is discussed in the text.

To examine how Noah-MP is improved from the baseline Noah LSM in terms of runoff simulation, we also present the comparison in Figure 2.6e, which indicates a

substantial improvement from the Noah LSM to Noah-MP. The results from the default Noah-MP setting slightly underestimate the USGS observations; however, they are already much better than the baseline Noah LSM results. The results from the calibrated Noah-MP are further improved, as both the magnitudes and the temporal variations correspond closely with the USGS observations. For easy comparison with similar studies, monthly climatological runoff is also shown in Figure 2.6d. Compared with previous studies by Falloon et al. (2011), Li et al. (2011), and Xia et al. (2012b), Noah-MP performs as well as or better than other mainstream LSMs in runoff modeling.

One must bear in mind that these improvements may be undermined by possible uncertainties in our comparison process and the models used. First of all, it is unfair to compare the model-simulated runoff directly (without river routing) with the USGS-gaged streamflow. However, the influence of the runoff routing on the comparison is relatively minor if we compare them at the monthly scale. Second, the USGS-gaged runoff is a direct measurement of the streamflow through a specific location without tracking its exact movement and distribution; therefore, great uncertainties evolve from human activities such as irrigation, tile drainage (Li et al., 2010), water supply, and reservoir regulation, as the MRB is one of the river basins that involve intensive water consumption (Murray et al., 2011). The traditional method is to use gaged streamflow to retrieve natural runoff without human interference. However, streamflow naturalization requires significant data on water use and water resource management, which are difficult to collect. Third, Noah-MP does not include a process to represent the artificial tile drainage in the model, which is a very important mechanism in some of Midwest areas, not only in runoff generation but also in groundwater and soil moisture modeling (Goswami, 2006; Algoazany et al., 2007; Gentry et al., 2009; Li et al., 2010). One possible solution is to improve the ability of Noah-

MP to represent human activities (e.g., irrigation and tile drainage) in future model development.

2.5.2 Groundwater

Groundwater dynamics have attracted increasing attention within the climate community (Fan et al., 2007; Miguez-Macho et al., 2007; Niu et al., 2007; Lo et al., 2010; Leung et al., 2011) for three reasons. First, groundwater directly influences soil moisture, which is an important variable in LSMs and climate models, an important indicator for drought detection, and a major controlling factor for the interaction between the land and the atmosphere (Niu et al., 2007). Second, groundwater, which provides most of the water needed for ET during the dry season (Gutowski et al., 2002), also influences ET. Because ET is both a water flux and a heat flux term, the influence of groundwater is passed on throughout the surface energy and water balances. Third, the inclusion of groundwater dynamics in climate models provides a direct tool to evaluate the impact of climate change on groundwater systems, which is vital for research into climate change adaptation.

Groundwater dynamics is one of the major improvements in Noah-MP; however, the employed SIMGM groundwater model was evaluated against the GRACE TWS anomaly data at the global scale (Niu et al., 2007) without comparison with actual ground measurements. Therefore, this study compares the simulated WTD against the USGS-observed WTD, both spatially and temporally. Figure 2.7 shows that Noah-MP can simulate the climatological spatial pattern, in which the water table is shallower in the southeast and deeper in the northwest. Some small areas in the wet region with deep water tables are not well simulated by Noah-MP, which may be due to the coarse spatial resolution or the model structure. Temporal variation is also compared with the observed groundwater storage data (Figure 2.8). The simulated groundwater variations agree very

well with the observations for the entire MRB, the Ohio-Tennessee River Basin, and the Upper MRB, with R^2 values of 0.75, 0.67, and 0.57, respectively. For the Ohio-Tennessee River Basin, the simulated anomalies are very similar to the observations. Because precipitation occurs frequently in this wet region, a small but very frequent fluctuation occurs in the observations, which Noah-MP fails to replicate. For the entire MRB and for the Upper MRB, the simulated anomalies are less than in the observations. For the Missouri River Basin, however, the simulated anomalies are much less than in the observations. The observed strong seasonal cycle is likely caused by the very shallow water table in this region (see Figure 2.7), which is because the aquifers are thin valley alluvium perched on top of the bedrock cores of the Rocky Mountains. These thin alluvial aquifers have very little storage, and thus, they are very responsive to seasonal snowmelt recharge (rises quickly) followed by efficient drainage into the deeply incised streams (i.e., the water level falls quickly). Models have difficulties representing these perched thin aquifers (Y. Fan, personal communication, 2011). For future model development, it would be helpful to collect the bedrock distribution data and include this process in LSMs.
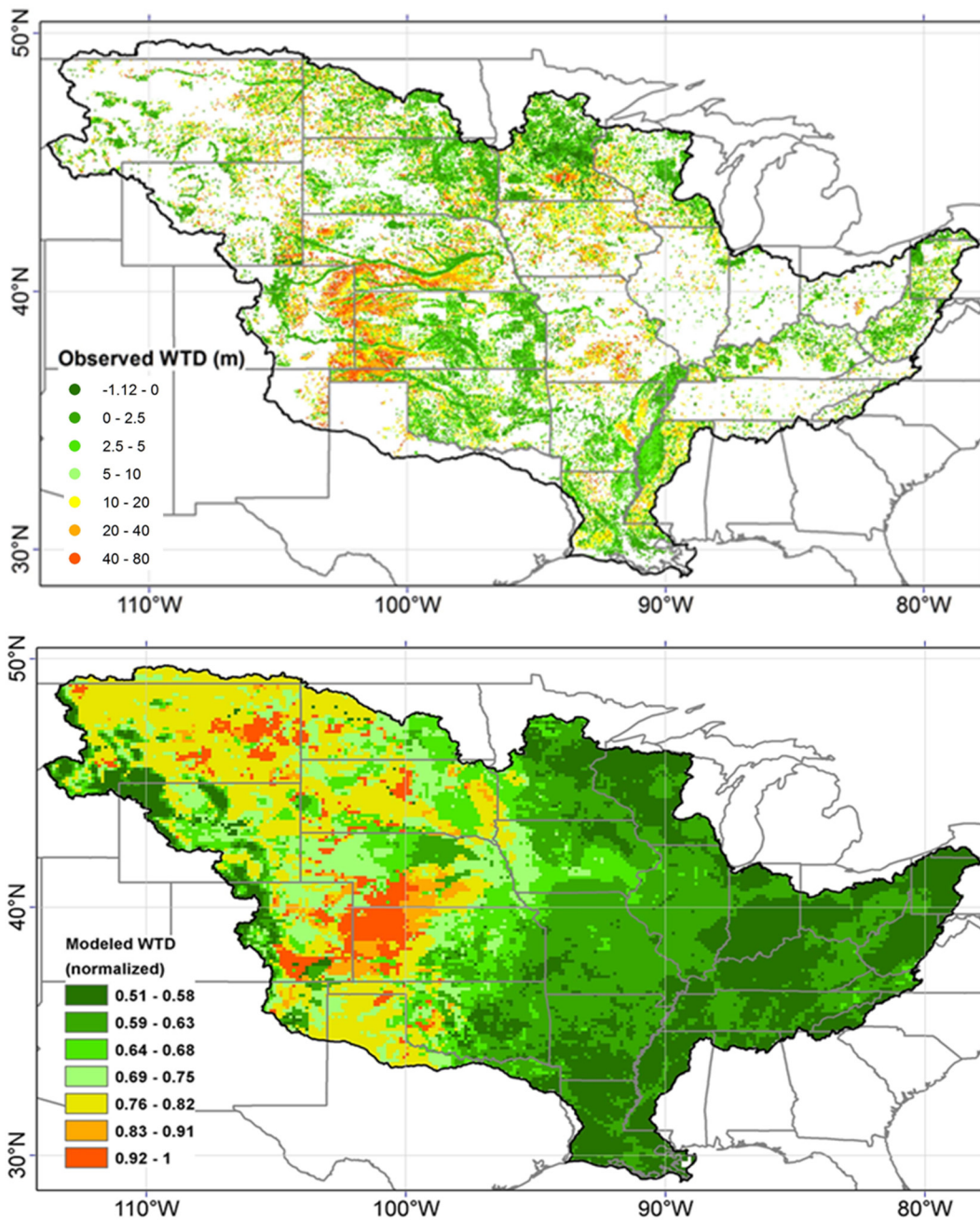
Figure 2.7 Climatological water table depth from (a) USGS measurements (all available observational time periods (Fan and Miguez-Macho, 2011)) and (b) Normalized Noah-MP simulations (2000–2009).
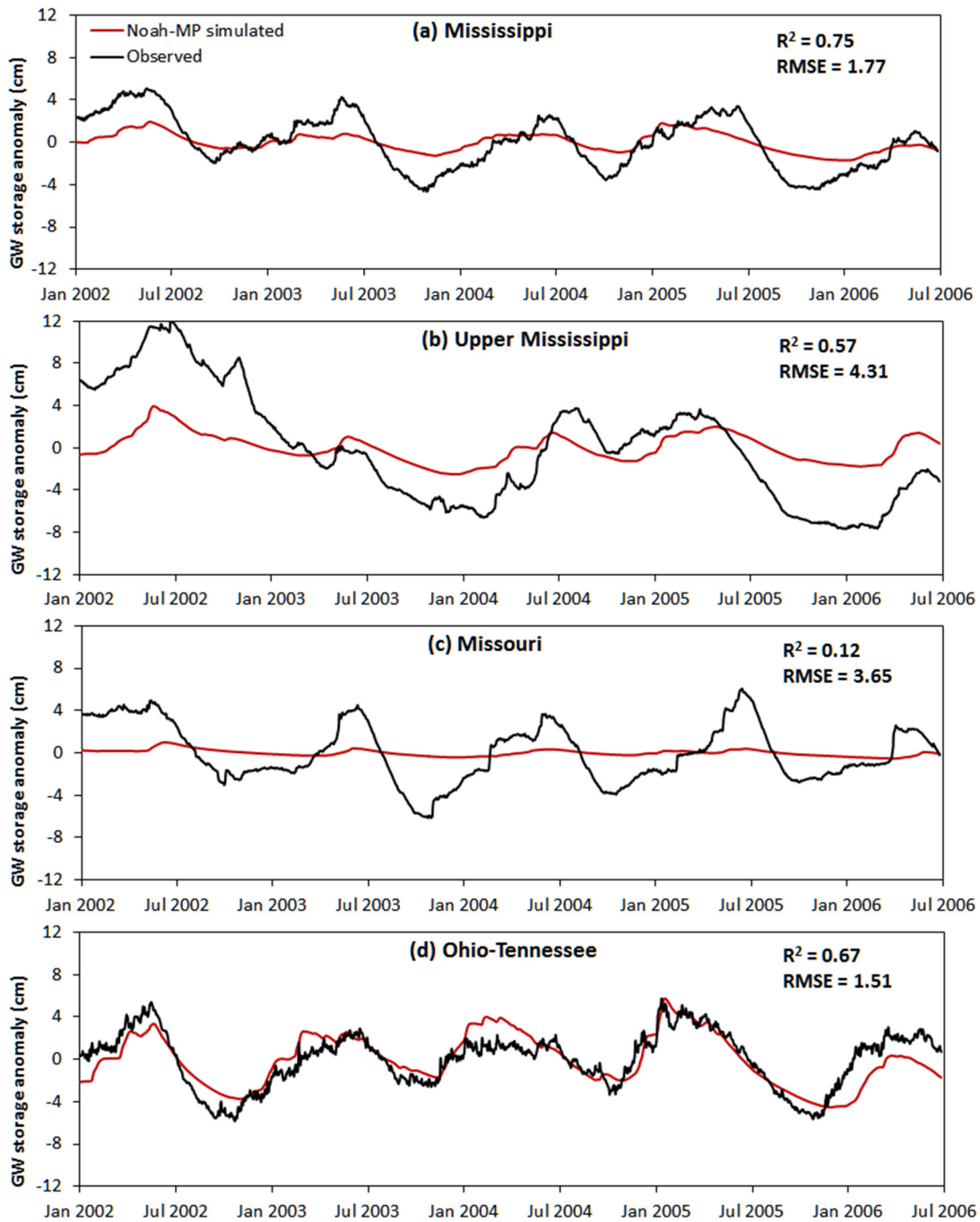
Figure 2.8 Comparison of the groundwater storage anomaly from observations (Rodell et al., 2007) and model simulations for (a) the entire Mississippi River Basin, (b) the Upper Mississippi River Basin, (c) the Missouri River Basin, and (d) the Ohio-Tennessee River Basin.

The modeled WTD is normalized here so that its spatial pattern is comparable to the observations. Indeed, Noah-MP-simulated WTD only ranges from 2 m to less than 14 m, whereas the observed WTD ranges from above the ground to greater than 80 m. The reason that WTD is limited to greater than 2 m in the model is to avoid a numerical computation problem. The simulated WTD cannot go deeper than 14 m, most likely due to the coarse spatial resolution. The range of WTD is very sensitive to the grid resolution: The finer the grid, the larger the range of WTD because when the grids are finer, the steeper land slope can be represented in the model, which accelerates the drainage speed. The most prominent scale for groundwater divergence-convergence is from hilltops to valleys. When averaging over the valleys and hills, and thus only having regional gradients, we get regional WTD and groundwater flow, which have much smaller gradients and ranges (Y. Fan, personal communication, 2011). Therefore, improving the spatial resolution is another direction for groundwater model development in LSMs, but there is always a tradeoff between model resolution and computational affordability.

2.5.3 Evapotranspiration

For the entire MRB, the simulated canopy evaporation, transpiration, and soil evaporation are 35.6 mm, 278.6 mm, and 323.5 mm, respectively, which account for 5.6%, 43.7%, and 50.7% of the total ET, respectively. To distinguish the effect of different vegetation types on ET, the 15 AmeriFlux tower sites are divided into four groups by their land cover types, with five grassland sites, five cropland sites, four forestland sites, and one shrubland site. Their climatological latent heat fluxes (the energy form of ET) from observations and from model simulations are presented in Figure 2.9. Among the four land cover types, forestlands and shrublands show significant improvements from Noah LSM to Noah-MP, in terms of better timing and more similar mean values, and grasslands also

38

have improved timing. However, we find that the Noah-MP-simulated latent heat fluxes are slightly higher than the observations from AmeriFlux for all four land cover types, which is similar to the evaluation by Blyth et al. (2010), whereas Noah LSM underestimates latent heat fluxes for forestlands and shrublands, overestimates for croplands, and well estimates the mean value for grasslands. Interestingly, the three land cover types for which Noah-MP performs well are grassland, forestland, and shrubland, which are considered more naturally occurring, whereas the land cover type for which Noah-MP does not perform well is cropland, which involves more human activities. This is most likely due to the use of leaf dynamics in Noah-MP, which can capture the processes of natural growth but cannot capture anthropogenic crop growth; thus, its ET increases too quickly during spring. In contrast, Noah LSM uses prescribed monthly LAI for various vegetation types and monthly GVF climatological values derived from NOAA/AVHRR, which better match anthropogenic crop growth. In addition to maintaining the strength in modeling natural vegetation dynamics, improvement in the simulation of the dynamic leaf model for cropland is highly recommended (this land type is expected to expand with the increasing population). One of the limitations of this study is that only runoff is calibrated; however, it is recommended that both runoff and ET be calibrated at the same time.
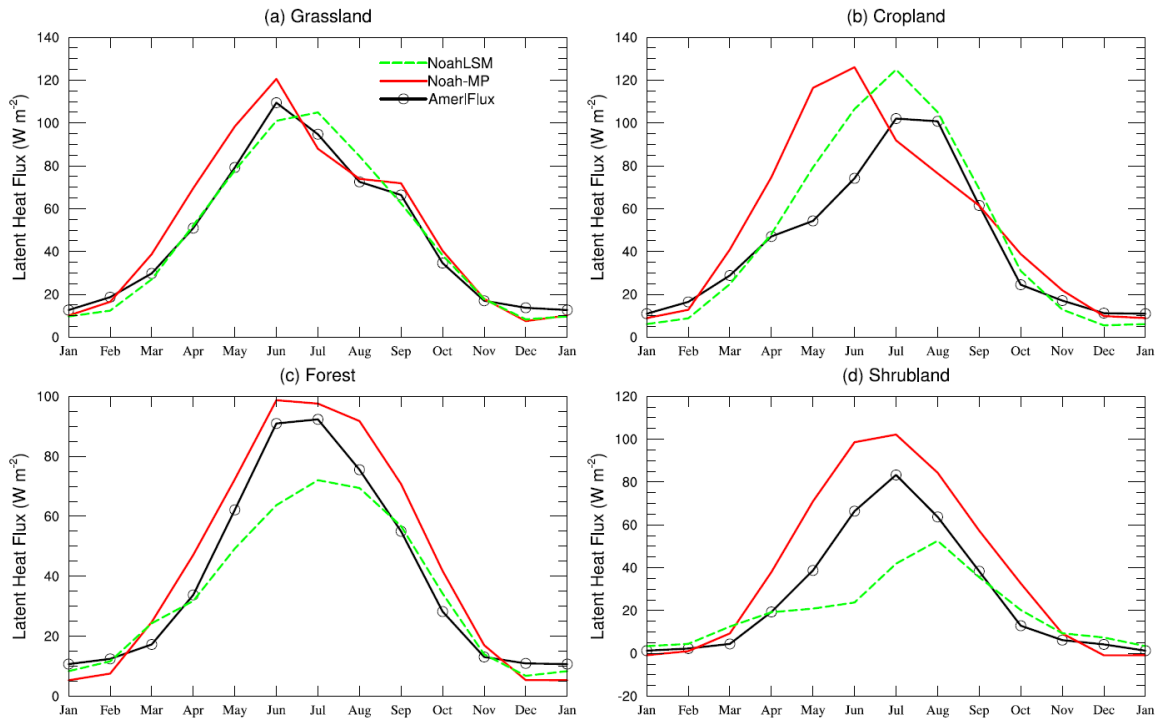
Figure 2.9 Comparison of the latent heat flux for the AmeriFlux observations and the model simulations for different land cover types. (a) Grassland (5 sites), (b) Cropland (5 sites), (c) Forestland (4 sites), and (d) Shrubland (1 site).

2.5.4 Soil Moisture

Studies (e.g., Entin et al., 2000) have shown that soil moisture measured at one location can represent the temporal variation for the surrounding area, up to 500 km in radius. Therefore, it is reasonable to use station-measured soil moisture to evaluate model-simulated soil moisture. Figure 2.10 compares the Noah-MP-simulated and SCAN-observed soil moistures for individual soil layers (the top 10 cm, 10–40 cm, 40–100 cm, and 100–200 cm). For the top layer (the top 10 cm), the Noah-MP-simulated soil moisture values are nearly identical to the SCAN observations, with an R^2 of 0.923 and an RMSE of 0.016. Because the top layer plays an important role in the water and energy exchanges between the land surface and the atmosphere, Noah-MP shows its high potential ability to study land-atmosphere coupling. For the second layer (10–40 cm), although the

40

discrepancy is slightly greater than the top layer (RMSE is 0.025), the comparison has an even higher R^2 value (0.933). A larger discrepancy is found in the deep layers, with R^2 of 0.624 and RMSE of 0.077 for the third layer (40–100 cm) and R^2 of 0.574 and RMSE of 0.035 for the bottom layer (100–200 cm). In terms of R^2, the results for the two deep layers are still acceptable; however, the Noah-MP-simulated soil moisture values largely underestimate the SCAN observations for the third layer, particularly in the summer and fall.

Why does Noah-MP underestimate the SCAN soil moisture for the third soil layer? This question can be answered by the connection with the ET comparison in the previous section, where it was demonstrated that Noah-MP-simulated ET values increased more quickly than did the AmeriFlux observations in the spring (Figure 2.9). Due to the high values of the simulated ET, more water is extracted from soil, which very likely leads to the low values of the simulated soil moisture. Furthermore, the variation in ET is dominated by transpiration (approximately twice that of soil evaporation). The transpiration rate from each soil layer is determined by the soil moisture factor controlling stomatal resistance, β_i (the higher value, the larger fraction of water for transpiration from the layer), which is parameterized as

$$\beta_i = \frac{\Delta z_i}{z_{root}} \min\left(1.0, \frac{\theta_{liq,i} - \theta_{wilt}}{\theta_{ref} - \theta_{wilt}}\right) \bigg/ \sum_{i=1}^{N_{root}} \left(\frac{\Delta z_i}{z_{root}} \min\left(1.0, \frac{\theta_{liq,i} - \theta_{wilt}}{\theta_{ref} - \theta_{wilt}}\right)\right) \qquad (2.10)$$

where Δz_i is the thickness of the ith soil layer (m); z_{root} is the total depth of root zone (m); $\theta_{liq,i}$ is the liquid soil moisture in the ith soil layer (m^3 m^{-3}); θ_{wilt} is the soil moisture at wilting point (m^3 m^{-3}); and θ_{ref} is the reference soil moisture (close to field capacity) (m^3 m^{-3}). In our comparison, most of the model grids with good availability of SCAN data (Figure 2.2) are croplands, grasslands, and shrublands, which have shallow roots, so that

root depth may only reach the third soil layer; only a few model grids with poor availability of SCAN data are forestlands, which have deep root depths that reach the bottom layer. Because of the great thickness of the third soil layer (0.6 m), the water supply for transpiration is heavily from the third layer. However, this may not be true in reality for croplands and grasslands, where roots may only reach the upper portion of the third layer not the entire third layer. Therefore, the water extraction for transpiration from the third layer is overestimated.

Figure 2.10c shows the strong annual cycle of transpiration, where high transpiration rates correspond to low soil moisture and low transpiration rates correspond to high soil moisture in the third soil layer. There are approximately 2 months of phasing difference between the transpiration and soil moisture, which is because after its peak, transpiration remains high and continues to dry the soil. To improve the soil moisture estimation for the third layer, adjustments (parameters or parameterization) to the dynamic leaf model in Noah-MP are needed to limit the increase in the ET rate in the spring.
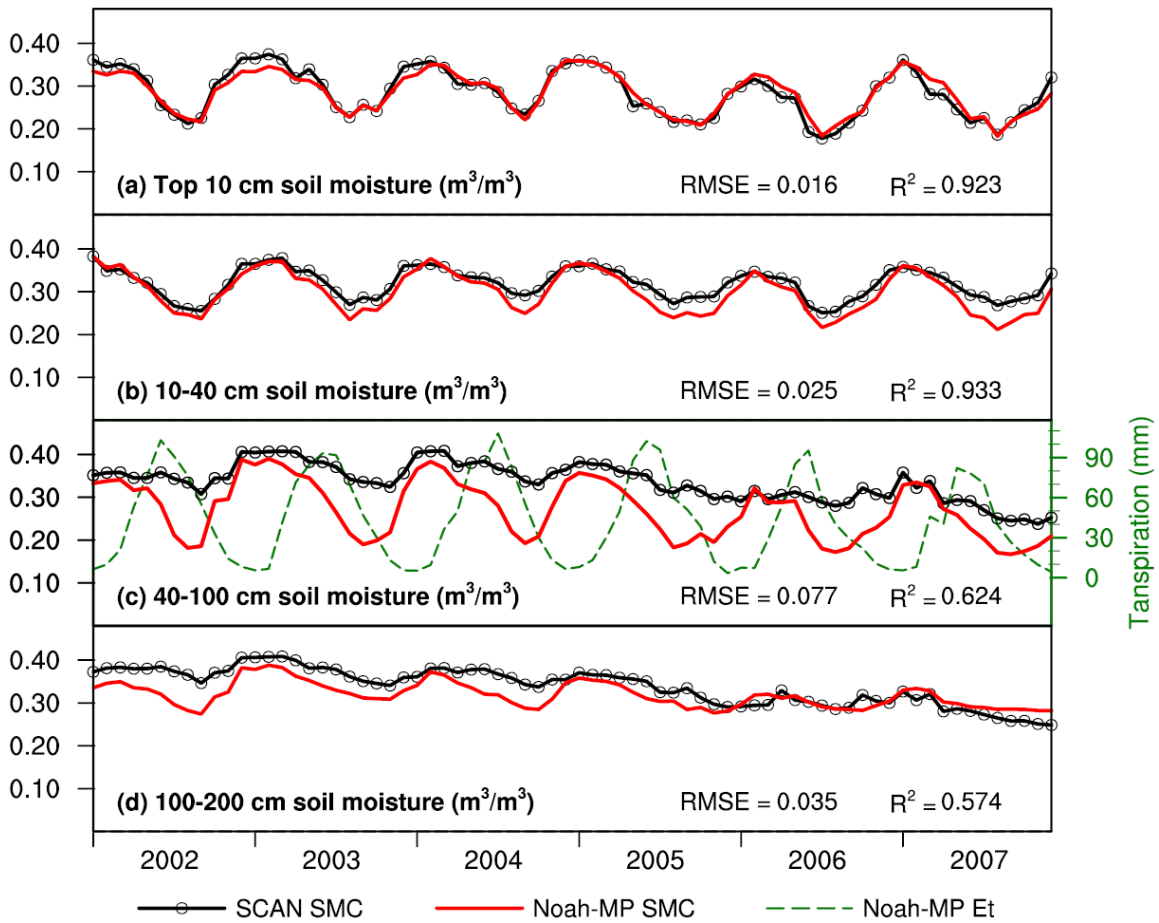
Figure 2.10 SCAN-observed and Noah-MP-simulated monthly soil moisture (SMC) for the the Mississippi River Basin at a depth of (a) top 10 cm, (b) 10–40 cm, (c) 40–100 cm, and (d) 100–200 cm. (c) also shows the Noah-MP simulated transpiration (Et). For the Noah-MP simulation, only those grids with a SCAN site are included, and for each grid, only those months with observed values are used.

2.5.5 Terrestrial Water Storage

The Noah-MP-simulated TWS anomaly is compared with the GRACE-based TWS anomaly for the entire MRB in Figure 2.11, in which the contributing components of the simulated TWS anomaly are also presented. There are several notable points from this comparison.

43

1. Noah-MP agrees well with GRACE in terms of the TWS anomaly, indicating that Noah-MP can capture the overall water cycle, including both the timing and the magnitude of water fluctuation. Although it may still involve great uncertainties from each of these components, Noah-MP captures the most important components such as soil moisture, groundwater and snow.

2. Because Noah-MP does not simulate the water storage of ice, lakes, rivers and biomass, from the water balance point of view, TWS has to be balanced by soil moisture, groundwater and snow, which are the water storage terms that are simulated in the model. In this particular region, soil moisture contributes the most to the TWS anomaly, followed by groundwater, and then snow. Although Noah-MP has difficulty in capturing the absolute values of WTD (see section 2.5.2), it is quite capable of capturing the annual groundwater fluctuation. Compared with the original Noah LSM, in which TWS is only balanced by soil moisture and snow, Noah-MP obtains a great improvement by including the second largest component of the TWS anomaly—groundwater.

3. Noah-MP still simulates only the natural part of the TWS anomaly, without considering human activities; therefore, Noah-MP has difficulty reflecting human interference. We can clearly observe that in the GRACE-based TWS anomaly curve, where there are two peaks in approximately half of the years, which cannot be observed in the Noah-MP curve. This is very likely due to human activities, for example, high irrigation rates in the spring. Some researchers have attempted to include irrigation in LSMs (Ozdogan et al., 2010; Pokhrel et al., 2012; Sorooshian et al., 2012).

4. In Figure 2.11b, the variation amplitude of the cumulative anomaly of ET is much higher than in the other fluxes (precipitation and runoff); hence, ET is the dominant water flux driving the TWS anomaly.
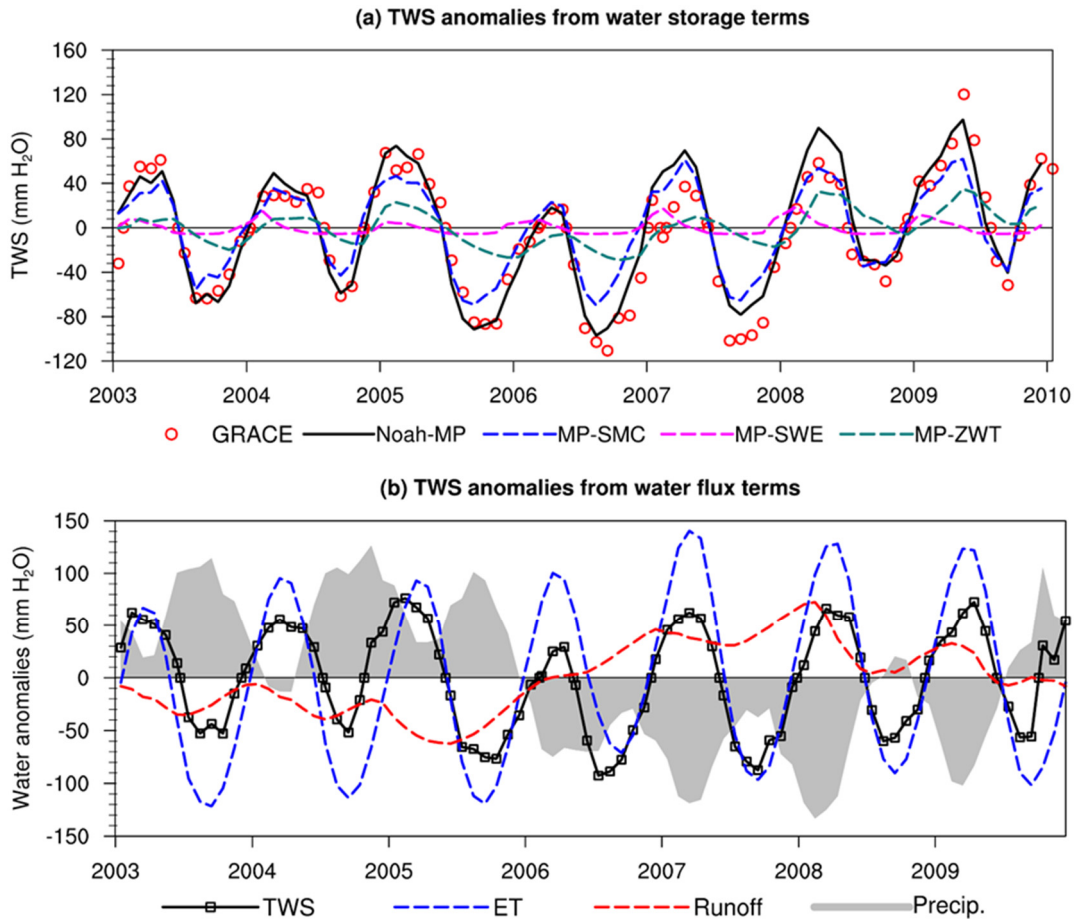
Figure 2.11 TWS anomalies for the Mississippi River Basin calculated from (a) the water storage terms and their contributing components and (b) the water flux terms and their contributing components. TWS anomalies are the cumulative anomalies of (Precip. – ET – Runoff), which are compared concurrently with the anomalies of the individual terms – ET and −Runoff. Note that the ET and runoff anomalies are shown as the negative of the original anomalies.

As shown in Figure 2.12, we also compare the Noah-MP-simulated TWS anomaly with the GRACE-based TWS anomaly at the subbasin level. In all four regions, soil moisture is always the largest contributor to the TWS anomaly; groundwater is the second largest contributor in the Ohio-Tennessee, Upper Mississippi and Lower Mississippi regions, but in the cold Missouri region, snow contributes as much as groundwater to the TWS anomaly. In the smaller subbasins, such as the Upper and Lower Mississippi, the

45

agreement between the model and the GRACE TWS anomalies is not as good as the agreement at the level of the entire MRB.
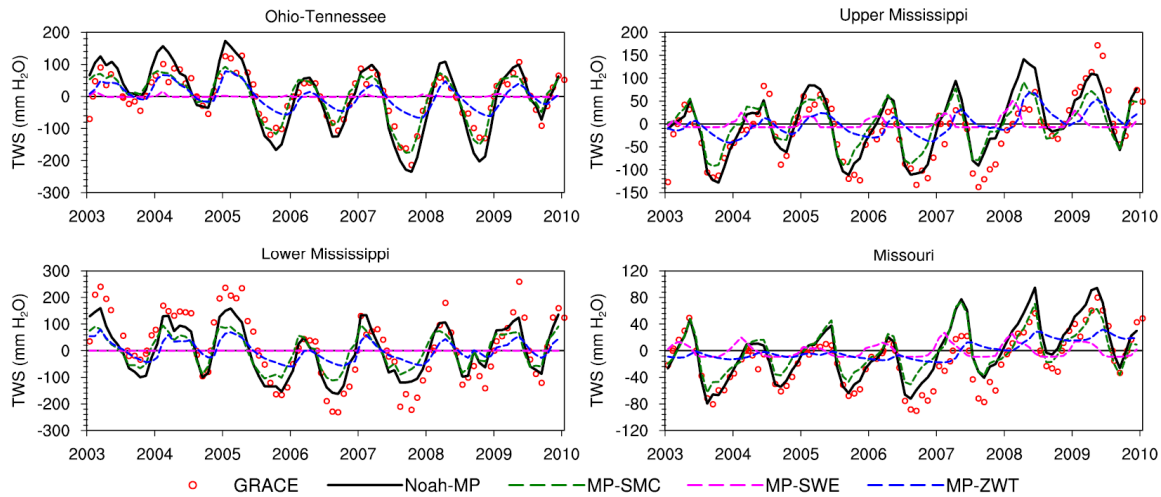
Figure 2.12 Comparison of the TWS anomalies from the GRACE-based measurements and the Noah-MP simulations from the water storage terms and their contributing components for the four subbasins.

2.6 CONCLUSIONS

In line with the GEWEX GLASS for LSM benchmarking or better model evaluation, we evaluated the model at the continental basin scale, specifically for the MRB. We began our evaluation with model spin-up, parameter sensitivity tests and model calibration, and then the calibrated results were compared with a number of traditional and recently available observational data sets. From this study, we have reached several conclusions that may be of interest to LSM developers and users.

With groundwater dynamics included in Noah-MP, it takes longer for WTD to reach equilibrium than without groundwater dynamics. This long WTD spin-up time would influence the spin-up times of other variables because when the water table is far from an equilibrium state, other variables such as runoff, ET, and soil moisture, need to be adjusted to help WTD reach equilibrium. For the entire MRB, at least 34 years is required for the

model to spin-up. For some mountain regions with very deep water tables, hundreds of years may be required for the model to spin-up.

Runoff is found to be sensitive to three parameters: the surface dryness factor (α), the saturated hydraulic conductivity (k), and the saturated soil moisture (θ_{max}); these three factors are selected for model calibration to improve runoff simulation. Although lumped calibration can improve model performance, distributed calibration is needed to obtain the best parameter values for some wet regions. If time and resources are limited for conducting automatic calibration (e.g., in this study), a better understanding of model physics and more analyses of the observational data would shorten the calibration time and benefit the model performance.

Noah-MP has shown significant improvements in hydrological modeling.

1. The Noah-MP simulated runoff is significantly improved compared with the baseline Noah LSM output in the NLDAS-2 framework. The spatial pattern of the Noah-MP simulated runoff matches fairly well with both the UNH-GRDC runoff and the USGS hydrologic unit runoff. We believe that this is the first time the USGS hydrologic unit runoff has been used in LSM evaluation and found to be very reasonable.

2. Groundwater evaluation indicated that Noah-MP captures the general spatial pattern of the climate conditions and captures the temporal patterns for the wet regions. However, it fails in simulating the absolute values and the temporal variation in the water table for the dry regions.

3. The addition of leaf dynamics to Noah-MP has improved its performance in ET simulation for natural land cover types.

4. One of the highlights of the study is that Noah-MP produces soil moisture values consistent with the SCAN observations for the top two soil layers (0–10 cm and 10–40 cm), which indicates its great potential for use in studying land–atmosphere coupling.

5. The Noah-MP-simulated TWS anomaly agrees very well with the GRACE observations, which may partly benefit from the inclusion in the model of groundwater dynamics, considered the second largest component of the TWS anomaly for most of the MRB.

2.7 ACKNOWLEDGEMENTS

Chapter 3: Assessment of Simulated Water Balance from Noah, Noah-MP, CLM, and VIC over CONUS Using the NLDAS Test Bed[2]

3.1 ABSTRACT

This study assesses the hydrologic performance of four land surface models (LSMs) for the conterminous United States using the North American Land Data Assimilation System (NLDAS) test bed. The four LSMs are the baseline community Noah LSM (Noah, version 2.8), the Variable Infiltration Capacity (VIC, version 4.0.5) model, the substantially augmented Noah LSM with multi-parameterization options (hence Noah-MP), and the Community Land Model version 4 (CLM4). All four models are driven by the same NLDAS-2 atmospheric forcing. Modeled terrestrial water storage (TWS), streamflow, evapotranspiration (ET), and soil moisture are compared with each other and evaluated against the identical observations. Relative to Noah, the other three models offer significant improvements in simulating TWS and streamflow and moderate improvements in simulating ET and soil moisture. Noah-MP provides the best performance in simulating soil moisture and is among the best in simulating TWS, CLM4 shows the best performance in simulating ET, and VIC ranks the highest in the simulations of streamflow. Despite these improvements, CLM4, Noah-MP, and VIC exhibit deficiencies, such as the low variability of soil moisture in CLM4, the fast growth of spring ET in Noah-MP, and the constant overestimation of ET in VIC.

3.2 INTRODUCTION

As a key component of weather/climate models, land surface models (LSMs) play an important role in weather prediction and climate projections (Pitman, 2003). Rapid

---

[2]Originally published as: Cai, X., Z.-L. Yang, Y. Xia, M. Huang, H. Wei, L.R. Leung and M.B. Ek, 2014b: Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed. *J. Geophys. Res.*, **119**(24): 13751-13770, doi:10.1002/2014jd022113. The References section contains full citations for all articles referenced here.

growth in LSM development has resulted in both the improvement of existing process representation and the addition of new processes and functionalities. However, it is a challenge to evaluate the accuracy of energy, water, and nutrient fluxes simulated by LSMs. Better model evaluation or land model benchmarking is the focus of several international projects. For example, the Global Energy and Water Cycle Exchanges Project (GEWEX) Global Land/Atmosphere System Study has identified LSM benchmarking as one of its three core activities (van den Hurk et al., 2011). The most recent activities are the diurnal land-atmosphere coupling experiment and the PALS (Protocol for the Analysis of Land Surface models) Land Surface Model Evaluation Benchmarking Project, with their descriptions available in the May 2013 and November 2013 issues of the GEWEX Newsletters, respectively. In addition, the International Land Model Benchmarking (ILAMB) project team has been working extensively to define evaluation strategies, identify benchmarks, create metrics, and improve model structure (Luo et al., 2012). In line with these efforts, a few existing platforms or tools are ready for the land model community to use, such as the PALS (Abramowitz, 2012) and the Land surface Verification Toolkit (Kumar et al., 2012).

The North American Land Data Assimilation System (NLDAS), which was originally designed to provide reliable initial land surface states to coupled weather/climate models (Mitchell et al., 2004), is another system that has the potential to be used for LSM benchmarking. Over the past decade or so, the NLDAS team has not only created high-quality atmospheric forcing data to drive LSMs but also collected substantial observational data sets to build necessary tools for evaluating the accuracy of surface and subsurface energy/water fluxes that LSMs produce. These efforts have resulted in the NLDAS test bed (Xia et al., 2013), which includes the following four LSMs: the community Noah LSM (Noah) (Ek et al., 2003), the Mosaic LSM (Mosaic) (Koster and Suarez, 1996), the

Sacramento Soil Moisture Accounting model (Burnash et al., 1973), and the Variable Infiltration Capacity (VIC) model (Liang et al., 1994). However, the current system does not yet include those models that incorporate recent developments in the land model community, such as the Community Land Model version 4 (CLM4) (Lawrence et al., 2011) and the multi-parameterization options version of the Noah model (Noah-MP) (Niu et al., 2011). Because these models feature improved physics and new functionalities, they can potentially provide better or new products that are not included in the current system. Therefore, the first motivation of this study is to examine the feasibility of using the NLDAS test bed to evaluate advancements in CLM4 and Noah-MP.

CLM4 and Noah-MP have adopted new model processes developed over the past decade or so, including interactive vegetation canopy, groundwater, and multilayer snow. As opposed to prescribed leaf area index (LAI) from observations, interactive vegetation canopy means that LAI is a prognostic variable that responds to the variability of precipitation, temperature, radiation, and nutrients availability (Dickinson et al., 1998; Thornton et al., 2007). Although introducing this type of dynamic leaf models may sometimes degrade model performance (Rosero et al., 2009; 2010), this approach adds vegetation as a memory process to the land system for seasonal climate forecasts (Jiang et al., 2009). In this study, both Noah-MP and CLM4 include dynamic leaf models that are from Dickinson et al. (1998) and Thornton et al. (2007), respectively. Another important improvement in both models is their inclusion of groundwater dynamics, thereby allowing the water table to modulate soil moisture and runoff. Several studies have shown the necessity of incorporating groundwater dynamics in offline or coupled LSMs (Fan et al., 2007; Miguez-Macho et al., 2007; Niu et al., 2007; Lo et al., 2010; Leung et al., 2011). Finally, Noah-MP (Niu et al., 2011) and CLM4 (Oleson et al., 2008) divide the snowpack into three and five layers, respectively, depending on total snow depth. Such multilayer

51

snowpack physics could potentially improve the accumulation and melt of snow and thus improve the timing of runoff generation. The addition of these processes has improved the models performance in simulating the hydrological cycle.

Additionally, this study takes advantage of the recent progress in making observational data available for model assessment. Alongside the high-quality small basin streamflow observation used in the previous NLDAS papers (Lohmann et al., 2004; Xia et al., 2012b), this study features the use of the Gravity Recovery and Climate Experiment (GRACE) satellite-derived terrestrial water storage (TWS) anomaly data, both the Moderate-Resolution Imaging Spectroradiometer (MODIS)-derived and the global network of micrometeorological flux tower sites (FLUXNET)-derived gridded evapotranspiration (ET) data, and the Soil Climate Analysis Network (SCAN)-observed soil moisture data. Instead of evaluating the streamflow only, a model's performance in simulating the major hydrological variables (TWS, streamflow, ET, and soil moisture) is comprehensively evaluated against these data sets. This is important not only for model development but also for facilitating the application of LSMs in hydrological, agricultural, and environmental studies.

In line with NLDAS-2 configuration, all four LSMs are driven by the same hourly NLDAS-2 atmospheric forcing, from 1979 to 2007, for the conterminous U.S. (CONUS) at 1/8° spatial resolution. Section 2 describes the atmospheric forcing data, the four LSMs and their configurations, and the model evaluation criteria. Section 3 introduces the observed and satellite-retrieved data sets. Section 4 shows the specific evaluations of TWS, streamflow, ET, and soil moisture. Section 5 summarizes the study.

3.3 METHODOLOGY

3.3.1 Models

Four LSMs are included in this study as mentioned in the introduction: Noah, VIC, Noah-MP, and CLM4. Table 3.1 shows a simple comparison of the four models. The Noah model is used here as the baseline LSM. The VIC model represents the "good" hydrological model, as it shows good performance in hydrological modeling in the model intercomparison by Xia et al. (2012b). Because Noah-MP and CLM4 incorporate recent major advancements, they are referred to as "advanced" LSMs in this study.

Table 3.1 Comparison of Noah, VIC, Noah-MP, and CLM4 in the treatments of vegetation, soil, and snow.

Model	Vegetation	Soil	Snow
Noah	Dominant vegetation type in one grid cell with prescribed LAI	4 layer moisture and temperature	Single layer
VIC	Tiling in one grid cell with prescribed LAI	3 layer moisture and temperature	2 layers
Noah-MP	Dominant vegetation type in one grid cell with dynamic LAI	4 layer moisture and temperature	Up to 3 layers
CLM4	Up to 10 vegetation types in one grid cell with prescribed LAI	10 layer moisture and 15 layer temperature	Up to 5 layers

Noah

The Noah model is the land component of both the community Weather Research and Forecast model (Skamarock et al., 2008) and the NOAA/NCEP (National Oceanic and Atmospheric Administration/National Centers for Environmental Prediction) weather/climate forecasting systems such as the Climate Forecast System and the Global Forecast System. Based on Noah version 2.7.1 used in the NLDAS-1 (Phase 1) (Mitchell et al., 2004), Noah version 2.8 is developed to improve snowpack simulation such as snow

water equivalent (SWE), snowmelt, and snow cover by adding snow age and an intermediate fix suggested by Slater et al. (2007). This fix reduces sublimation and increases SWE for all stable cases including snow-free grid cells (Livneh et al., 2010). It also improves energy fluxes, streamflow, and land surface temperature simulation for warm seasons (Smith et al., 2012; Wei et al., 2013). This version of Noah is used to generate NLDAS-2 Noah products using a quasi-operational mode (Xia et al., 2013). NLDAS-2 Noah products have been comprehensively evaluated using *in situ* observations (Xia et al., 2012b; 2013; 2014a; 2014b).

VIC

The VIC model is a macroscale, semidistributed, grid-based, hydrologic model developed at the University of Washington and Princeton University (Liang et al., 1994; Wood et al., 1997). VIC can be executed in several modes; the full water and energy balance version was chosen for NLDAS-1 (version 4.0.3) and NLDAS-2 research work (version 4.0.5). The model includes three soil layers, with a 10 cm top layer and two deeper layers of spatially varying thicknesses. The root zone can span all three layers, depending on the vegetation type and its associated vertical root distribution. Like the Mosaic model, the VIC model utilizes subgrid vegetation tiles. The VIC model includes a two-layer energy balance snow model (Cherkauer et al., 2003) that not only represents snow accumulation and ablation on the ground and in the forest canopy but also uses sub-grid elevation bands to represent the impact of elevation on temperature, precipitation, and snow. VIC has been widely applied to large river basins all over the world (Nijssen et al., 1997; Lohmann et al., 1998), at national (Maurer et al., 2002; Livneh et al., 2013) and global scales (Sheffield et al., 2006). In addition, NLDAS-2 VIC products (version 4.0.5) have been evaluated through the NLDAS test bed framework (Xia et al., 2012b; 2013; 2014b).

54

Noah-MP

Building on the Noah model (Ek et al., 2003), Noah-MP (Niu et al., 2011; Yang et al., 2011) incorporates the recently improved physics in LSMs. It also employs the idea of using a single model for ensemble forecasting. Major improvements implemented in Noah-MP include the following: (1) modifying the model structure to include a one-layer canopy and three-layer snow, (2) adding a semi-tile subgrid scheme that separates vegetated areas from bare ground to better account for the surface energy balance, (3) introducing a more permeable frozen soil by separating permeable and impermeable fractions, (4) developing a TOPMODEL-based runoff scheme and a simple groundwater model to improve the modeling of soil hydrology, and (5) adding a short-term leaf dynamic model to simulate LAI and vegetation greenness fraction. Noah-MP features multiparameterization options for dynamic leaf, canopy stomatal resistance, runoff and groundwater, a soil moisture factor controlling stomatal resistance (the β factor), and six other processes. Each combination of these scheme options forms an ensemble member for ensemble forecasting. Therefore, instead of using multiple models, a single land surface model—Noah-MP—can be used for ensemble forecasting. The scheme options used in the study are: the modified Dickinson et al. (1998) scheme for leaf dynamics, the Ball-Berry scheme for canopy stomatal resistance, the Noah-type soil moisture factor controlling stomatal resistance, the TOPMODEL runoff with the SIMGM groundwater, the Monin-Obukhov scheme for surface exchange coefficient for heat, the Niu and Yang (2006) scheme for both supercooled liquid water (or ice fraction) and frozen soil permeability, the modified two-stream scheme for radiation transfer, the Canadian Land Surface Scheme-type scheme for ground snow surface albedo, and the Jordan (1991) scheme for partitioning precipitation into rainfall and snowfall.

CLM

CLM4 is the land component of the Community Earth System Model (formerly known as the Community Climate System Model) (Gent et al., 2010; Lawrence et al., 2012). Compared to its previous versions, CLM4 was enhanced with various representations of hydrological processes, including those associated with runoff generation, groundwater dynamics, soil hydrology, snow module, and surface albedo (Lawrence et al., 2011). Specifically, runoff generation is parameterized using a simplified TOPMODEL-based representation (Niu et al., 2005; Oleson et al., 2010). The 42 m deep ground is divided into 10 hydrologically active layers (i.e., the "soil" layers) extending from the surface to 3.8 m and five bottom thermal layers that are not hydrologically active to accurately capture soil temperature dynamics in century-scale integrations (Alexeev et al., 2007; Nicolsky et al., 2007). The thermal and hydrologic properties of organic soil are considered based on Lawrence and Slater (2008). Soil water is then calculated using a revised numerical solution of the one-dimensional Richards equation (Zeng and Decker, 2009). Recharge to groundwater from the soil column and groundwater table depth is updated dynamically following the algorithm described in Niu et al. (2007). The snow module is updated with new snow cover and snow burial fraction parameterizations (Niu and Yang, 2006; Wang and Zeng, 2009), as well as with the Snow and Ice Aerosol Radiation model, to describe grain size-dependent snow aging and vertically resolved snowpack heating (Flanner et al., 2007). In addition to the above-mentioned processes, the model has been extended to include a carbon-nitrogen biogeochemical model that is capable of simulating vegetation, litter, soil carbon and nitrogen, and vegetation phenology prognostically (Oleson et al., 2010).

CLM4 considers spatial heterogeneity of the land surface using a nested subgrid hierarchy—a grid cell is composed of a number of land units including glacier, lake,

56

wetland, urban, and vegetated surfaces. The latter could then be decomposed further to 15 possible Plant Functional Types (PFTs) plus bare ground. In this study, the percentages of land cover types and PFTs within each CLM4 grid cell are derived from the MODIS-based global land parameter data set in Ke et al. (2012) at 0.05° resolution. The model is configured to be driven by satellite phenology as (1) the prognostic biogeochemical and vegetation phenology modules are not activated (dynamic vegetation is turned off) and (2) vegetation phenology (i.e., leaf and stem area indices) is prescribed using the data set from Ke et al. (2012). Soil texture is generated based on a hybrid of 30 arc second State Soil Geographic Database (now referred to as the U.S. General Soil Map) (Miller and White, 1998). The two-layer soil type data are then converted to a composition of clay and sand (Cosby et al., 1984) within each 30 arc-second grid cell and interpolated to 10 vertical layers down to a 3.8m depth.

3.3.2 Forcing Data

All four LSMs use the same atmospheric forcing data from NLDAS-2 (Xia et al., 2012c). The seven forcing fields are precipitation, air temperature, specific humidity, surface air pressure, wind speed, incoming solar radiation, and incoming longwave radiation. NLDAS-2 atmospheric forcing data are derived from North American Regional Reanalysis (NARR) data and supplemented with several sources of observed data sets such as the unified gauge-based precipitation from NOAA Climate Prediction Center and the satellite-derived downward shortwave radiation. The coarser NARR data, at the 32 km spatial resolution and 3 hour temporal resolution, is interpolated to the finer NLDAS data at 0.125° resolution and at a 1 hour time interval (Cosgrove et al., 2003a). In line with the NLDAS-2 atmospheric forcing, all LSMs are run at 0.125° latitude-longitude resolution for the CONUS and at hourly time step.

3.3.3 Model Spin-Up and Integration Procedure

Given certain atmospheric forcing, static inputs (e.g., soil and vegetation types), and parameters, individual LSMs needs to find their own equilibrium state before models can be used for "real" simulations and applications. Yang et al. (1995) discussed the spin-up time required for a variety of 22 LSMs over two sites (tropical forest and grassland, respectively). Results showed that, normally, models took less than 12 years to reach equilibrium. However, for those LSMs with groundwater dynamics, the required spin-up time is much longer (Niu et al., 2007; Cai et al., 2014a).

In the NLDAS-2 project, to minimize the impact of model initialization on model products, a two-step method was used to generate the initial states for Noah and VIC (Xia et al., 2012c). First, the last 8 years of output from the 11 year NLDAS-1 real-time simulation from 1 October 1996 to 31 December 2007 were averaged for each model to produce climatological initial land states for 00Z 1 October. In the second step, a spin-up simulation was carried out as follows: (1) the climatological 00Z 1 October states from NLDAS-1 were used to initialize a 15.25 year spin-up simulation for each model from 1 October 1979 to 1 January 1995 using the NLDAS-2 forcing data; (2) the 1 January states from the last 10 years of this spin-up simulation were then averaged together to provide initial states for the final NLDAS-2 simulation.

In this study, Noah-MP is initialized by running the model repeatedly through the entire year of 1979 44 times (44 data years). CLM4 is spun-up by recycling the NLDAS-2 forcing in 1979–2007 for 36 cycles (i.e., ~1000 years) until all state variables, including soil moisture, temperature, and groundwater table depth, reached equilibrium. Such a long spin-up period is necessary due to the deep ground column of CLM4 compared to other models.

3.3.4 Model Output

After spin-up, Noah-MP and CLM4 were executed from 1 January 1979 to the present using the NARR- and observation-based NLDAS-2 forcings. Because the routing code that translates gridded model runoff to streamflow needs at least six months of spin-up, the first 9 month (1 January 1979 to 30 September 1979) run is used as a spin-up period and only 28 years of model outputs from 1 October 1979 to 1 September 2007 were analyzed in this study.

The model outputs from Noah and VIC are identical to those used in Xia et al. (2012c), while the model outputs from Noah-MP and CLM4 are newly run for this study. All four models provide the following outputs: runoff, evapotranspiration, soil moisture, and SWE. In addition, Noah-MP and CLM4 also provide groundwater water table depth (WTD) in order to calculate the TWS anomaly.

The TWS anomaly is calculated from storage terms that include two variables (soil moisture and SWE) for the Noah and VIC models and three variables (soil moisture, SWE, and WTD) for the Noah-MP and CLM4 models (Equation 3.1). Because there is no explicit representation of rivers, lakes, or reservoirs, water storage in these water bodies are not included. This exclusion is not realistic; however, it is reasonable because these water bodies are excluded from the calculation of water balance in LSMs as well.

$$TWSA_i = \begin{cases} SMCA_i + SWEA_i, & \text{for Noah and VIC} \\ SMCA_i + SWEA_i + WTDA_i, & \text{for Noah - MP and CLM4} \end{cases} \tag{3.1}$$

where $TWSA_i$, $SMSA_i$, $SWEA_i$, and $WTDA_i$ are the anomalies of TWS, soil moisture, SWE, and WTD, respectively.

3.3.5 Lumped Routing Model and Calculation of Streamflow

As streamflow is measured at gauging stations, for fair comparison, model simulated runoff at each grid cell must be routed to the basin outlet. To do so, the NLDAS test bed provides both a flow-direction mask for the CONUS and a routing model. The flow-direction mask was provided by the NOAA/NWS/OHD (National Oceanic and Atmospheric Administration/National Weather Service/Office of Hydrologic Development), which calculates the amount and timing of runoff both within each grid cell and on the river network, assuming that water in a grid cell can only go into one of its neighboring grid cells. The routing model can be operated in either lumped mode or distributed mode (Lohmann et al., 2004). The simple lumped mode is used in this study because (1) the lumped model has been calibrated for these small basins and has shown relative success compared to the distributed mode and (2) it is relatively easy to use. The routing model convolutes the sum of each model's runoff in each basin with one impulse response function $UH(t)$ (Lohmann et al., 2004; Xia et al., 2012b). This function is solved by deconvoluting

$$Q_{obs}(t) = \frac{\Delta\tau}{86.4}\left(\sum_{\tau=0}^{\tau_{max}} S^i R_{t-\tau}^i\right) UH_\tau \tag{3.2}$$

where Q_{obs} is the observed streamflow (m³/s), $\Delta\tau$ is the time interval of the streamflow observation (1 day), S^i is the area of a grid cell in the basin (km²), R^i is the modeled runoff of a grid cell (mm/d), 86.4 is the factor for unit conversion, and τ_{max} is the length of the impulse response function $UH(t)$. The term τ_{max} is less than 7 days for all study basins and represents the maximum concentration time of runoff within the basin.

3.3.6 Model Evaluation Criteria

To quantify the differences between model simulations and observations, the following evaluation statistics are used: spatial distribution of difference, relative bias, correlation coefficient (r), coefficient of determination (R^2), and root-mean-square error (RMSE). Relative bias (RB) is calculated as

$$RB = \frac{\overline{M} - \overline{O}}{\overline{O}} \tag{3.3}$$

where $\overline{M}$ is the mean of modeled streamflow and $\overline{O}$ is the mean of observed streamflow.

3.4 OBSERVED AND SATELLITE-RETRIEVED DATA

3.4.1 GRACE Terrestrial Water Storage

The GRACE twin satellites, which were launched in March 2002, measure the temporal change of the Earth's gravity field with time, based on the variation in the distances between the two satellites (Tapley et al., 2004). As other mass changes are relatively small, GRACE satellites primarily detect the changes in TWS. GRACE-derived TWS data have been widely used in studies on LSM development (Niu et al., 2007; Lo et al., 2010; Leng et al., 2013; Lei et al., 2014), groundwater depletion (Famiglietti et al., 2011), and drought detection/assessment (Houborg et al., 2012; Chen et al., 2013; Long et al., 2013). This study uses the GRACE TWS anomaly data, processed by Landerer and Swenson (2012), based on the CSR RL5.0 release from the Center for Space Research at the University of Texas at Austin. The data are at 1°×1° spatial resolution and monthly temporal resolution and can be freely downloaded from the Jet Propulsion Laboratory TELLUS website (http://grace.jpl.nasa.gov). This is very helpful to users who do not have the expertise in removing the errors in the GRACE TWS data.

3.4.2 U.S. Geological Survey Streamflow

U.S. Geological Survey (USGS) streamflow data gauged at 961 small basins are used for streamflow evaluation. The 961 small basins, ranging from 23 km^2 to 10,000 km^2, are selected from the original 1145 small basins used in the NLDAS-1 (Lohmann et al., 2004). The selection is based on expert judgment by the NOAA/NWS/OHD (Xia et al., 2012b), with the purpose of removing those small basins with visible signs of reservoir operation. However, we cannot remove all management effects. The total area of the 961 small basins is 1.57 million km^2, which is approximately 20.6% of the CONUS. These small basins have good representation over the eastern part (east of 95°W), covering about 50% of the land surface area; while the basin representation is poor over the middle and western parts. Monthly streamflow data for the period from 1 October 1979 to 30 September 2007 are converted from the daily streamflow obtained from the USGS website.

3.4.3 MODIS and FLUXNET Evapotranspiration

Two different measurements are used for ET evaluation: MODIS ET and FLUXNET-MTE (model tree ensemble) ET. The MODIS global ET products are from Mu et al. (2011), which have been improved by (1) simplifying the calculation of vegetation cover fraction, (2) calculating ET as the sum of daytime and nighttime components, (3) adding calculation of the soil heat flux, (4) improving estimates of stomatal conductance, aerodynamic resistance, and boundary layer resistance, (5) differentiating between dry and wet canopy surfaces, and (6) dividing the soil surface into saturated and moist components (Mu et al., 2011). Overall results show that the improved algorithm increases ET values globally and overcomes the negative bias that existed in previous ET products generated by Mu et al. (2007). Monthly and annual versions of these improved products were obtained from Mu at 0.125° resolution. The data cover the entire NLDAS-2 domain and span the period from January of 2000 to September of 2007.

The FLUXNET-MTE data integrates 253 FLUXNET eddy covariance towers distributed over the globe using a machine learning technique, the model tree ensemble (MTE) approach (Jung et al., 2009; 2010). The MTE approach processes the gap filled half hourly eddy covariance fluxes to global fluxes at 0.5°×0.5° spatial resolution and monthly temporal resolution. The data has been used in ET trend detection (Jung et al., 2010) and model improvement (Bonan et al., 2011). With the dense distribution of FLUXNET sites over the CONUS, we particularly expect high data quality for our study domain. The FLUXNET-MTE ET data are available for the whole period of 1982–2008; however, to keep consistent with the MODIS data, we only use the period of 2000–2007.

3.4.4 SCAN Soil Moisture

For evaluation of soil moisture simulations, we use soil moisture data from the Soil Climate Analysis Network (SCAN)—a nationwide soil moisture and climate information system led by U.S. Department of Agriculture/Natural Resources Conservation Service (Schaefer et al., 2007). Soil moisture data are collected by dielectric constant measuring devices at 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm, where possible. The data used in this study are identical to those used in Liu et al. (2011) and Xia et al. (2014b). By performing extensive quality control steps, the following invalid data have been removed: values outside a reasonable range, inconsistent data affected by sensor calibration or installation, and data measured under frozen conditions. Figure 3.1 shows the 121 SCAN stations in the CONUS and their data availability. To overcome the low data availability at most stations while distinguishing different climate conditions, the data are aggregated into six regional averages (Southeast, Northeast, Texas, Great Plains, Northwest, and Southwest, respectively, see Figure 3.1 for detail) as suggested by Xia et al. (2014b), on a monthly basis.
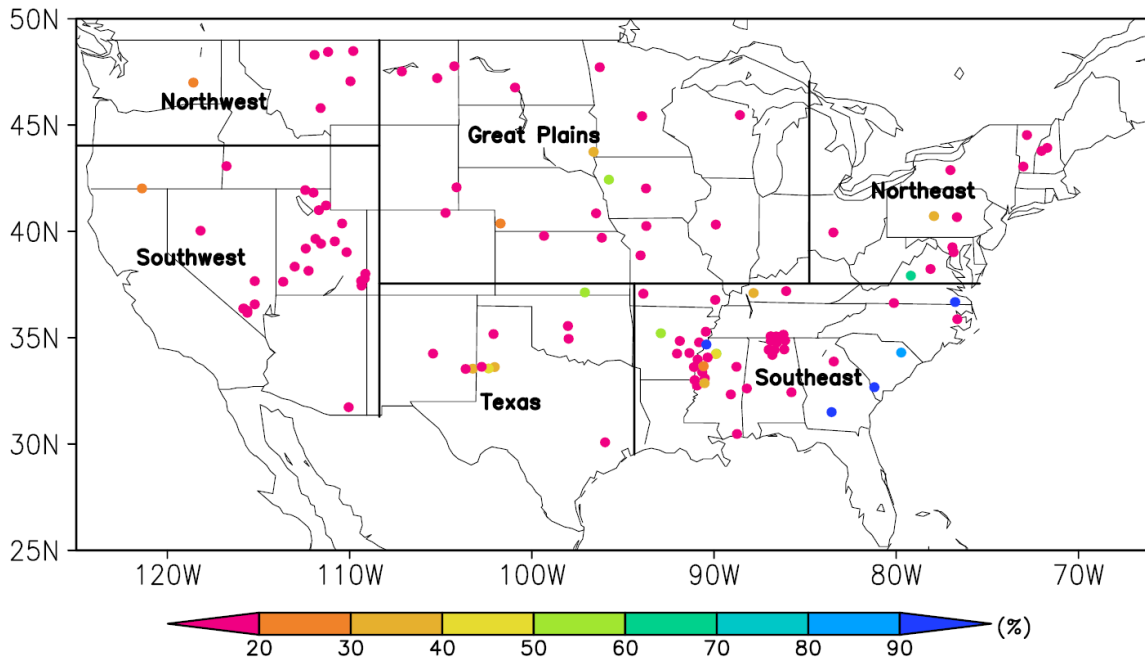
Figure 3.1 Spatial map showing the geographical locations and data availability of the 121 SCAN sites. Black solid lines dividing the conterminous United States into six regions: Southeast, Northeast, Texas, Great Plains, Northwest, and Southwest, respectively.

3.5 EVALUATION OF MODEL PRODUCTS

3.5.1 Evaluation of TWS Simulation

Fluctuation of TWS in a region reflects the departure of dry/wet conditions from its climatology. It involves the variation in soil moisture, groundwater, snow, ice, and water storages in lakes, rivers, and biomass. Therefore, the skill in modeling TWS can reflect the overall performance of a land model in hydrological simulation.

Figure 3.2 shows the amplitude of the annual TWS variation from GRACE and model simulations. To examine the effect of different climate conditions on the model performance, the CONUS is divided into four quadrants: Northwest, Northeast, Southwest, and Southeast, based on the 40°N latitude line and the 98°W longitude line (Figure 3.2a). The spatial pattern of the GRACE TWS variation generally follows the spatial pattern of

64

annual precipitation (e.g., the precipitation pattern shown in Cosgrove et al. (2003a)), with very high precipitation/TWS variation in the Northwestern region and the Southeastern region and very low precipitation/TWS variation in the Southwestern region. Noah-MP best captures this pattern and the other three LSMs can also reproduce this general pattern. However, Noah simulated TWS variation is too low in the Northwestern and Southeastern regions where GRACE TWS variation is high. Although CLM4 simulation generally follows this pattern, its simulated amplitudes are either too high in the Northwestern and Southeastern regions or too low in the Southwestern region (Figure 3.2d).

There are two things that GRACE and the LSMs do not agree on: (1) The GRACE-derived TWS variation along the coastal lines is extremely low, which is likely due to the fact that the GRACE signal is contaminated by the signal from ocean, while all LSMs do not show this pattern. (2) In the Southeastern quadrant, the deep blue line in the GRACE map (Figure 3.2a) is overlaid with the Mississippi River channel, which indicates the very high TWS variation amplitude due to the water fluctuation in the river. However, this cannot be clearly seen in the LSM maps (Figures 3.2b-e), which is probably because these LSMs do not explicitly consider change of water storage in rivers and the exchange of groundwater and river water in their formulations.
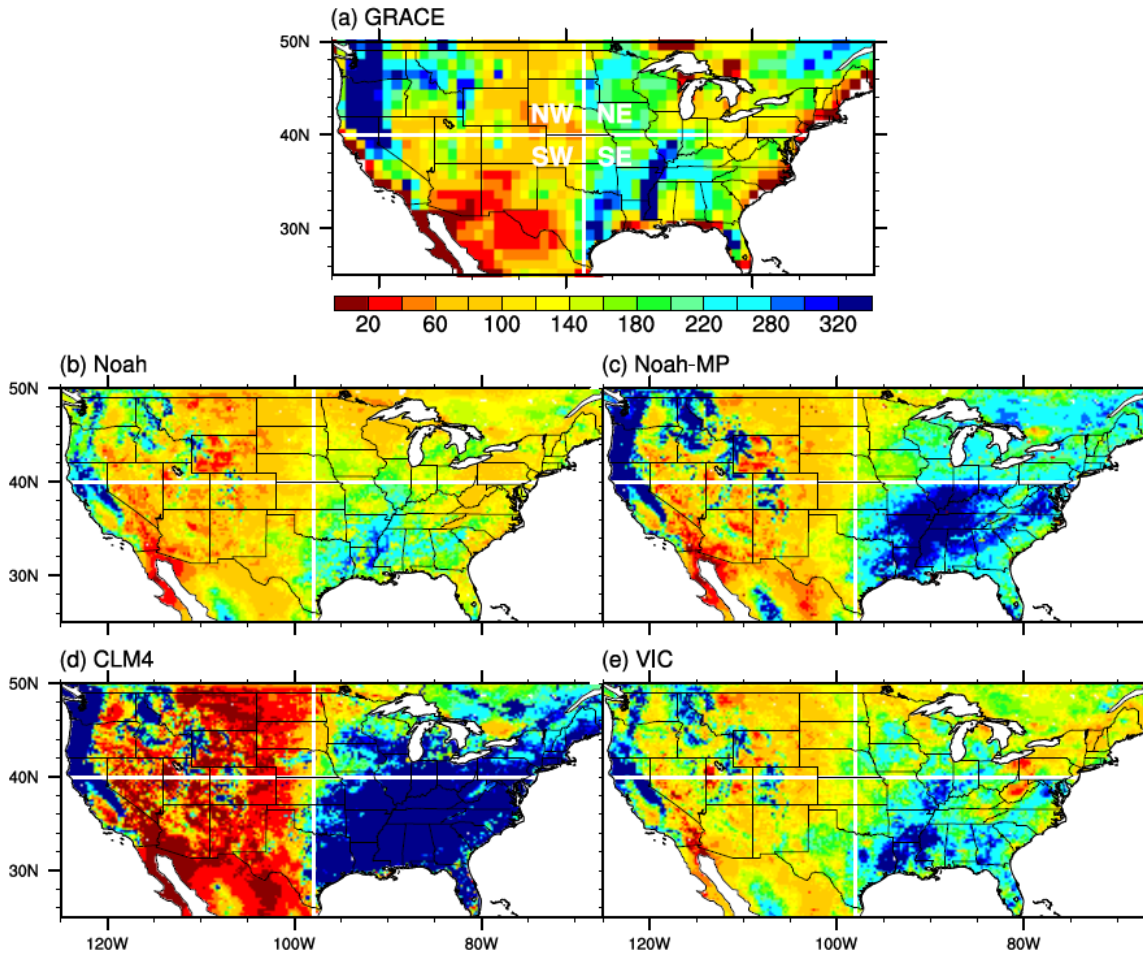
Figure 3.2 Amplitude of the annual TWS variation (2003–2007) from (a) GRACE, (b) Noah, (c) Noah-MP, (d) CLM4, and (e) VIC. Annual TWS variation is defined as the difference of the maximum and the minimum of the monthly TWS in a year. The white solid lines divide the CONUS into four quadrants, as indicated in (a). The spatial resolution is 1°×1° for GRACE and 0.125°×0.125° for all four LSMs.
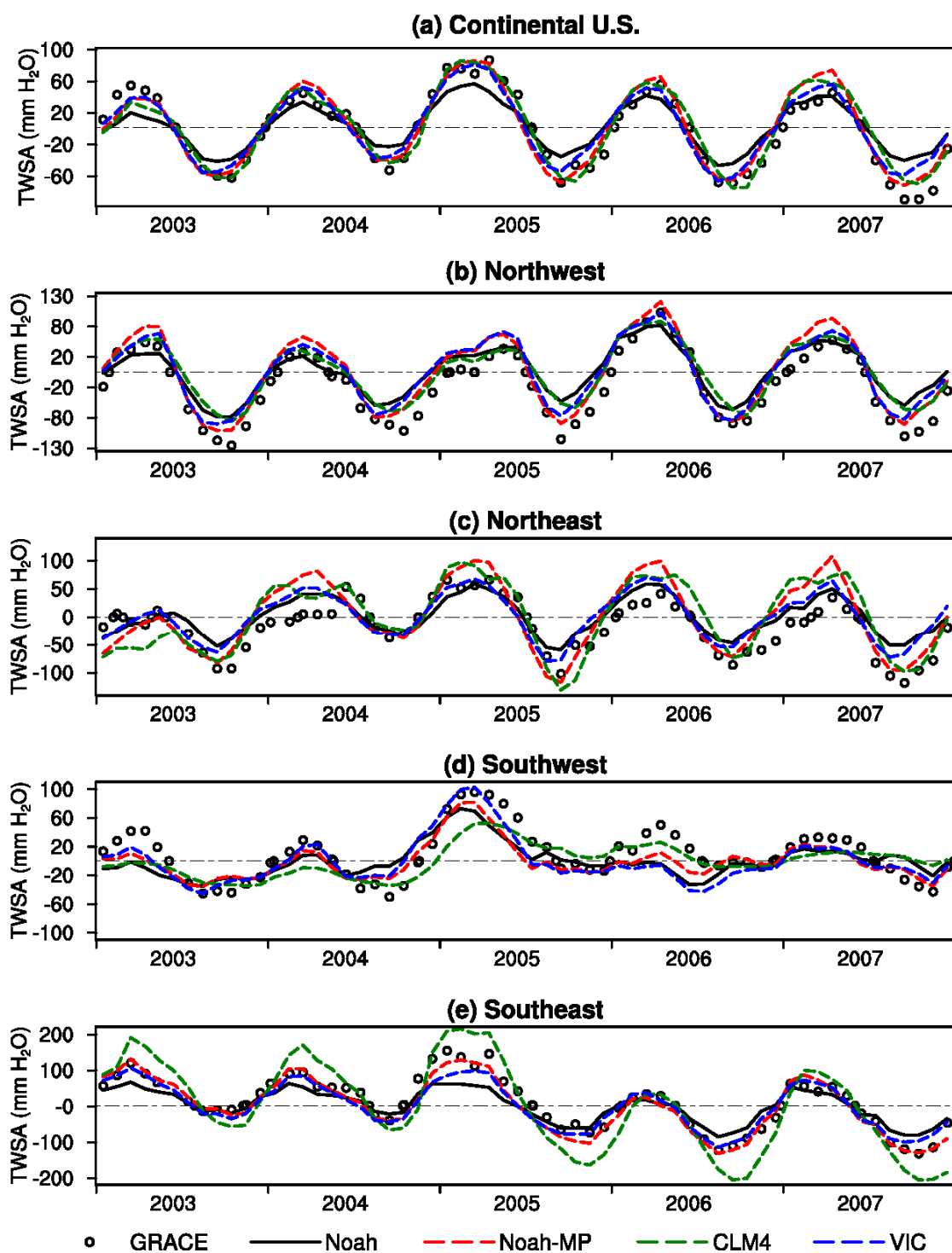
Figure 3.3 GRACE-derived and LSMs-simulated terrestrial water storage anomaly over the (a) Conterminous U.S., (b) Northwest, (c) Northeast, (d) Southwest, and (e) Southeast.

Figure 3.1 shows a comparison of TWS anomaly between the four model simulations and the GRACE observation, with their statistics of model performance summarized in Table 3.2. For the entire CONUS (Figure 3.1a), all four LSMs perform reasonably well. The largest discrepancy is between the Noah model and GRACE. Although Noah captures the timing, its amplitude is smaller than GRACE. The other three models capture the TWS fluctuation. Statistics show that R^2 ranges from 0.894 (Noah) to 0.913 (CLM4) and RMSE ranges from 14.50 (CLM4) to 22.55 (Noah). The improvement of the advanced LSMs in R^2 is minor (0.907–0.913 for the advanced LSMs compared to 0.894 for Noah), while the improvement in RMSE is large (both below 15.17 for the advanced LSMs compared to 22.55 for Noah).

Table 3.2 Statistical summary of model performance in simulating terrestrial water storage anomaly, based on comparison with GRACE observation. All R^2 values pass the 99% confidence level. The thick underlines indicate the highest R^2 or lowest RMSE among the four land surface models.

Model	R^2						RMSE					
	NW	NE	SW	SE	Avg.	CONUS	NW	NE	SW	SE	Avg.	CONUS
Noah	0.914	<u>0.739</u>	0.534	0.917	0.776	0.894	30.89	<u>30.41</u>	25.89	34.51	30.42	22.55
Noah-MP	<u>0.962</u>	0.696	<u>0.790</u>	<u>0.932</u>	<u>0.845</u>	0.907	<u>24.47</u>	38.05	<u>19.65</u>	<u>21.97</u>	<u>26.03</u>	15.17
CLM4	0.956	0.683	0.671	0.912	0.805	<u>0.913</u>	26.29	38.81	23.33	57.10	36.38	<u>14.50</u>
VIC	0.933	0.694	0.670	0.906	0.801	0.906	26.10	31.31	22.35	25.16	26.23	15.50
Mean	0.941	0.703	0.666	0.917	0.807	0.905	26.94	34.64	22.81	34.68	29.77	16.93

Model performance is evaluated in each of the four quadrants as well. In the Northwest quadrant (Figure 3.1b), all LSMs perform very well featuring very high R^2 and low RMSE, with the highest R^2 and lowest RMSE from Noah-MP, while the lowest R^2 and highest RMSE are from Noah. The reason all four LSMs perform well in this region is probably because the TWS anomaly signal is dominated by the snow process—particularly in the Rocky Mountains and the Northwestern Costal Mountains—which GRACE (the

LSMs) can easily detect (simulate) the signal. In the Northeast quadrant (Figure 3.1c), all model performances are also acceptable, but compared to the other three quadrants, models tend to produce low R^2 and high RMSE. Surprisingly, the highest R^2 and lowest RMSE are from Noah, while the lowest R^2 and highest RMSE are from CLM4. This is the only quadrant that advanced LSMs show no obvious improvement to Noah. The amplitudes of annual TWS variation simulated by Noah and VIC are smaller than those simulated by Noah-MP and CLM4, which can be more clearly seen in Figure 3.2 (less blue shaded area in Figures 3.2b and 3.2e than in Figures 3.2c and 3.2d). In the Southwest quadrant (Figure 3.1d), all LSMs feature relative low R^2 and low RMSE. The low RMSE by all four LSMs does not necessarily mean low discrepancy, rather it is more due to the low TWS anomaly variation in this region. As we can see from Figure 3.1d, except for the abnormal high TWS anomaly in 2005, the TWS anomaly is less than 60 mm in all other years, much smaller than the 100 mm for the entire CONUS. This low TWS anomaly variation can be seen more visually in Figure 3.2, particularly in the U.S. Southwest Border region. Similar to the Northwest region, Noah-MP produces the highest R^2 and lowest RMSE while Noah produces the lowest R^2 (0.534, the only one below 0.6) and highest RMSE. Lastly, in the Southeast quadrant (Figure 3.1e), CLM4 is able to capture the temporal variation as shown by the high R^2 (0.912). However, it produces very high RMSE (57.10) due to the much larger variation of TWS anomaly compared to GRACE (Figure 3.2d). The other three LSMs perform very well featuring very high R^2 and low RMSE, with the highest R^2 and lowest RMSE from Noah-MP. Similar to the Northwest region where the TWS anomaly signal is strong due to the high snowfall rate, the TWS anomaly signal is strong in this region too due to the high precipitation rate. Therefore, both GRACE and LSMs can easily detect these strong snow and precipitation signals.

We also summarize model performance one by one. The Noah model can simulate the general spatial and temporal patterns, but its simulated TWS variation is smaller than the GRACE observation. Overall, it produces lower R^2 and higher RMSE than the three advanced LSMs. Noah-MP's performance is interesting. It is not the "best" model in terms of high R^2 and low RMSE for the entire CONUS; however, if we average the R^2 and RMSE for each of the four quadrants, it produces the highest averaged R^2 (0.845) and lowest averaged RMSE (26.03). Because Noah-MP produces good results across all four quadrants of CONUS, it shows the most robust performance in this study. In contrast, although CLM4 shows the best statistics for the entire CONUS, its performances in different regions vary. In particular, the amplitudes of annual TWS anomaly variation are either too large (Eastern CONUS) or too small (Western CONUS), with very small transition areas. Lastly, VIC performs as robust as Noah-MP, with only small differences in statistics (R^2 and RMSE).

Although the Noah and VIC models do not include the groundwater dynamics, they can still capture the TWS anomaly. However, groundwater does make a significant contribution to the anomaly. As shown in Table 3.3, groundwater anomaly contributes about 36% to the R^2 and about 32% to the RMSE on average over the CONUS. The snow contribution to TWS anomaly is largely dependent on the location—large contribution over heavy snow regions and negligible over snow free regions; while contributions from soil moisture and groundwater are more stable among different regions.

Table 3.3 Contribution of soil moisture (SMC), snow water equivalent (SWE), and groundwater (GW) to the R^2 and RMSE between GRACE and model simulated TWS (%).

	NW			NE			SW			SE			CONUS		
	SMC	SWE	GW	SMC	SWE	GW	SMC	SWE	GW	SMC	SWE	GW	SMC	SWE	GW
R^2															
Noah-MP	44.0	29.6	26.4	48.4	21.2	30.4	33.4	43.8	22.8	50.7	4.2	45.2	42.1	24.4	33.5
CLM4	41.3	20.0	38.7	42.3	16.3	41.4	36.2	37.1	26.7	50.8	0.2	49.0	39.8	18.7	41.5
RMSE															
Noah-MP	25.2	34.2	40.7	28.7	36.5	34.8	29.0	35.4	35.6	19.9	47.2	33.0	23.1	38.0	38.9
CLM4	36.3	36.6	27.0	34.8	35.5	29.8	33.3	35.9	30.8	33.0	40.7	26.4	39.7	42.0	18.3

3.5.2 Evaluation of Streamflow Simulations

Streamflow simulations by the four LSMs are evaluated against USGS observed streamflow using relative bias and correlation coefficient. Figure 3.4 shows the relative bias of the mean annual streamflow for the four LSMs. The color shaded areas are the 961 USGS small river basins. We find that the LSMs have difficulty representing streamflow in the north central region (as indicated by the red box in Figure 3.4a). As suggested by Yang et al. (2011), Noah-MP overcomes Noah's large overestimation of streamflow mainly due to the use of the M-O scheme for the surface exchange coefficient (C_H), which results in a smaller C_H than using the scheme from Chen et al. (1997). In the large area centered on the southeastern part of the Midwest (the center is indicated by the blue circle in Figure 3.4a), all models show their best performance featuring low relative bias within 20% (grey shaded area). CLM4 shows the largest good performance area (almost the entire east of 95°W), followed by Noah-MP and VIC, with Noah the least (only the area bounded by the blue circle in Figure 3.4a). In the Pacific Northwest region, the Noah model overestimates streamflow for most areas with only some underestimates occurring in some small, isolated areas; while the other three LSMs present approximately equal amounts of overestimates and underestimates.
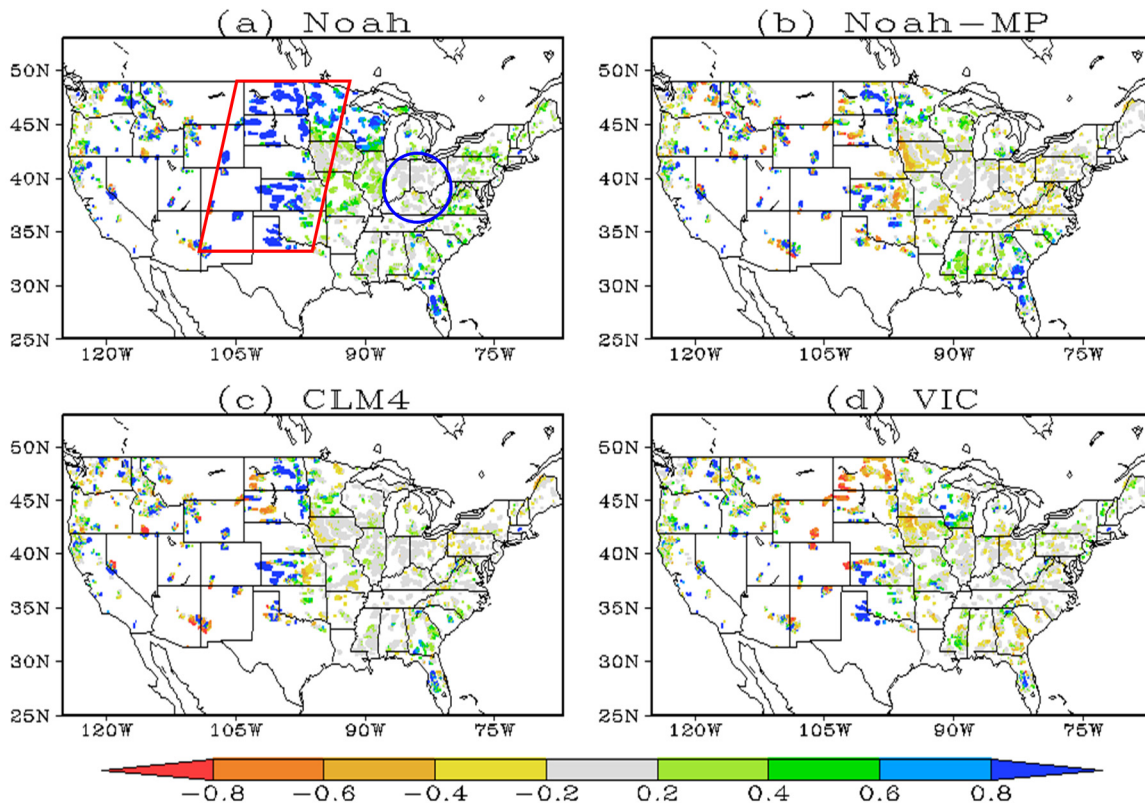
Figure 3.4 Relative bias of mean annual streamflow for the 961 small river basins from October 1979 to September 2007 from (a) Noah, (b) Noah-MP, (c) CLM4, and (d) VIC.

The relative bias presented here is generally much larger than we usually see in literature regarding modeled streamflow bias for a specific river basin (Christensen et al., 2004). For a single river basin (usually medium or large basins), we can optimize model parameters (calibration) to minimize the relative bias to a small fraction, e.g., ±0.05. However, it is a big challenge to calibrate a model to fit each of many small basins simultaneously, because some parameters are location sensitive (Hogue et al., 2006). Therefore one set of parameter values may not be suitable for every river basin. For example, a fixed set of hydrologic parameters were used over all grid cells in the CLM4 simulation, while previous studies (Hou et al., 2012; Huang et al., 2013) have demonstrated that CLM4 surface flux and runoff calculations are sensitive to its hydrologic parameters

72

over different watersheds. Furthermore, these biases can be attributed to uncertainties from atmospheric forcing data, static input data (e.g., soil and land cover types), or observation bias. For example, the streamflow simulation has been improved over the northern Rocky Mountains due to the adjustment of precipitation forcing with the monthly climatological precipitation data from the Parameter-Elevation Regressions on Independent Slopes Model (Xia et al., 2012b).

Figure 3.5 shows the correlation coefficients between the model simulated and the USGS observed streamflow for each of the small basins. The spatial pattern of these correlation coefficients is similar among all four LSMs, with the very high correlation values presented in the West Coast and Eastern U.S. (particularly the Southeastern U.S.) and low correlation values in the northern Midwest. Overall, the Noah (VIC) model shows low (high) correlation values for more small basins than the other three LSMs do; while the Noah-MP and CLM4 models present more medium high correlation values. The cause for the low correlation in Noah has been investigated by Xia et al. (2014a). The results show that after the strong constraint of aerodynamic conductance (Livneh et al., 2010) is released in snow-free grids, correlation values were largely increased, particularly in the Northeastern region. The higher correlation in VIC may be partly due to its calibration work (Troy et al., 2008) and partly due to its development philosophy from the hydrology community (Xia et al., 2012a). The improvement in Noah-MP and CLM4 is mainly attributed to the better simulation of snow and the buffering effect of groundwater dynamics. Particularly, Noah-MP produces very high correlation values over the Pacific Northwest snow region (see the red box in Figure 3.5b), which means Noah-MP better captures the seasonal cycle of snow.

73
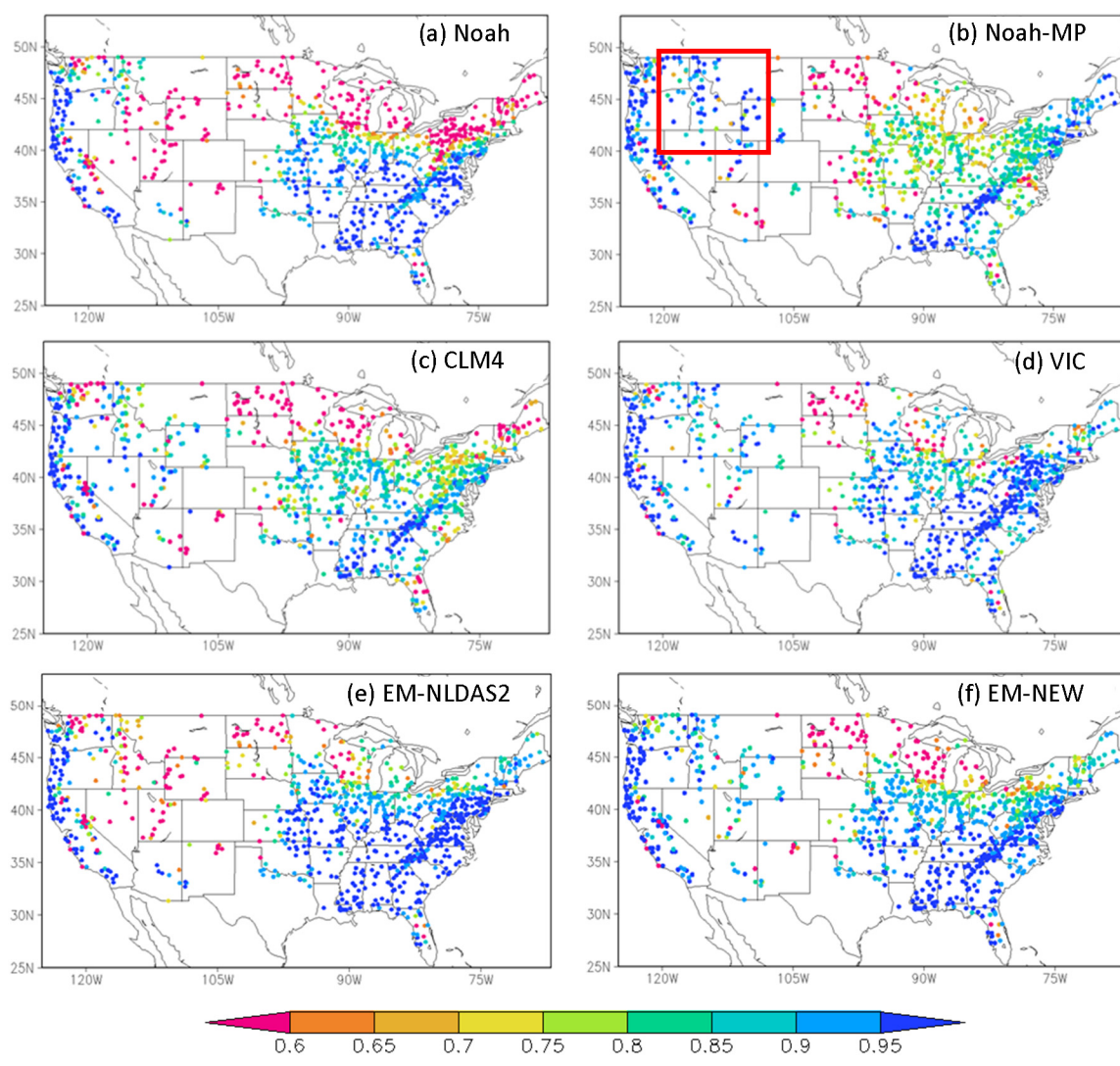
Figure 3.5 Correlation coefficient of the 28-year monthly climatological streamflow between model simulations and USGS observation for the 961 small river basins for (a) Noah, (b) Noah-MP, (c) CLM4, (d) VIC, (e) ensemble mean of the NLDAS-2, and (f) ensemble mean of this study.

74
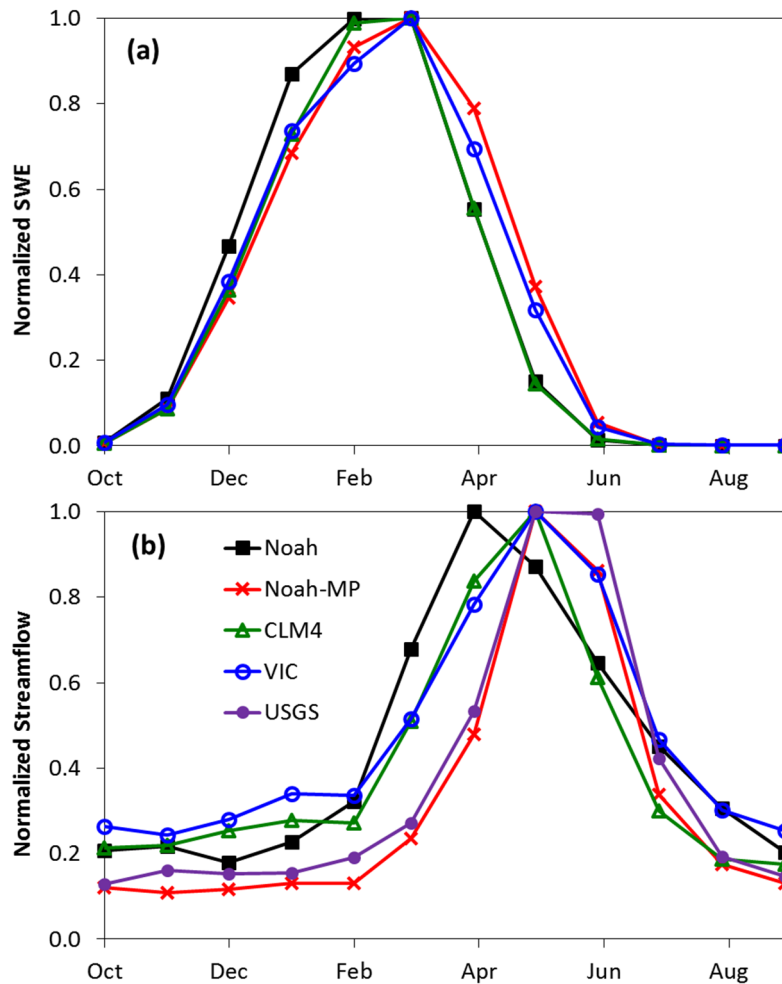
Figure 3.6 Model simulated and/or USGS gauged monthly climatology (averaged from October 1979 to September 2007) of (a) snow water equivalent (SWE) and (b) streamflow over the area highlighted by the red box in Figure 3.5b. To show the timing of snow accumulation and melting more clearly, we use the normalized SWE and streamflow.

To test this, Figure 3.6 shows the seasonal cycle of SWE and streamflow over this region. Comparing with the Noah model, we see a timing difference in Noah-MP as well as in VIC and CLM4. Previous studies revealed that Noah simulates snowmelt too early (Livneh et al., 2010). Our study also shows the early snowmelt in Noah, which leads to the early streamflow peak. In comparison, Noah-MP simulates snowmelt slowest so that the timing of the simulated streamflow matches best the USGS observation. These

75

improvements in Noah-MP are presumably due to their better representations of ground heat flux, retention, percolation, and refreezing of melted liquid water within the multilayer snowpack (Yang et al., 2011). Consequently, the ensemble mean of this study (Figure 3.5f) exhibits significant improvements over that of the current NLDAS-2 land models (Figure 3.5e).

3.5.3 Evaluation of ET Simulation

Figure 3.7 shows the comparison of climatological ET between model simulations and observations from both FLUXNET and MODIS. Similar to the evaluation of TWS anomalies, the CONUS is divided into four quadrants: Northwest, Northeast, Southwest, and Southeast, based on the 40°N latitude line and the 98°W longitude line. To quantify the uncertainties in observation, simulations are treated as "perfect" fit if they are within the two time series defined by FLUXNET and MODIS. Using this standard, CLM4 is good for all four quadrant regions. Noah is good for all four regions except for the Northeast and Great Lakes where ET is largely underestimated. The reason is that ET is strongly constrained by the aerodynamic conductance under stable boundary layer conditions applied over snow-free grid cells (Xia et al., 2014a). Noah-MP is good for the Southwest region only, while VIC simulated ET is constantly higher than the two observations.

We also find Noah-MP simulated ET grows very fast starting in February in the Northeast (Figure 3.7b) and Southeast (Figure 3.7d) regions, which results in the largest ET difference between Noah-MP simulation and the two observations in May. This is very likely due to the use of the dynamic leaf phenology model in Noah-MP, which predicts LAI as a function of light, temperature, and soil moisture, rather than using the prescribed monthly LAI as the Noah model does. This occurs in the two Eastern quadrant regions

because the vegetation effect on ET is higher in these two regions due to the higher percentage of forest coverage.
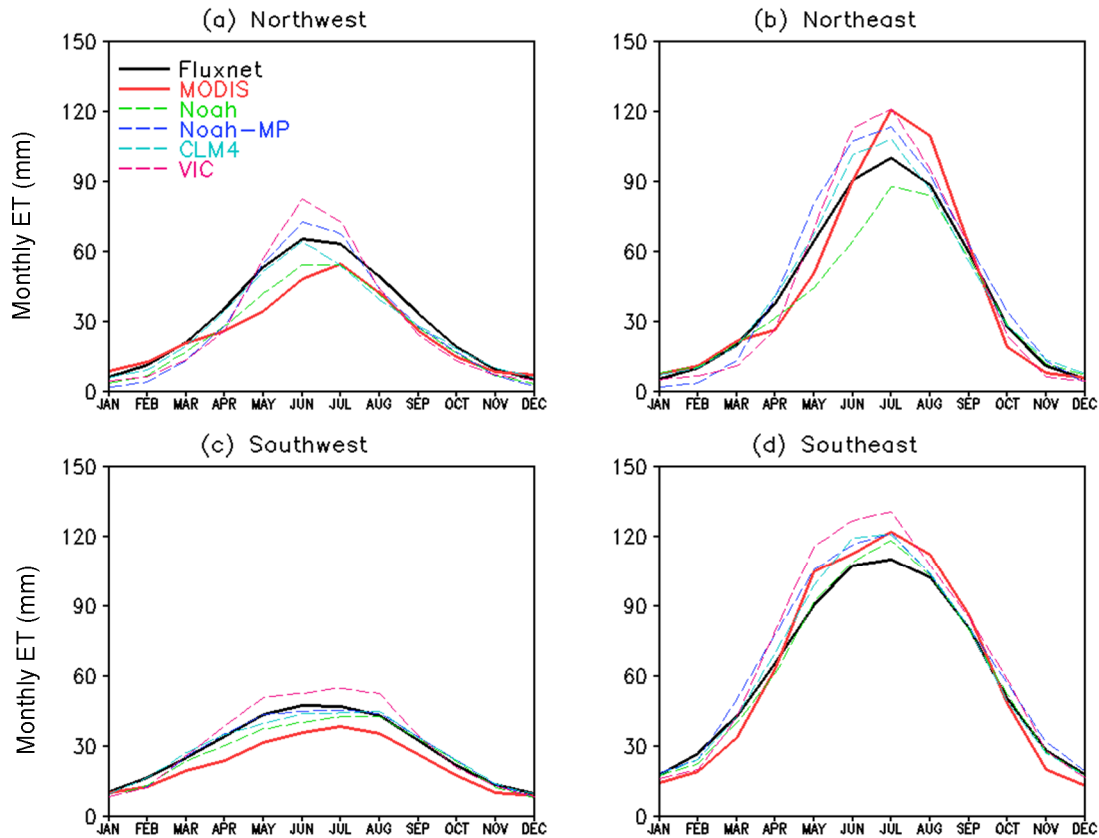
Figure 3.7 Comparison of monthly evapotranspiration (ET) climatology (2001–2007) between simulations and observations (both FLUXNET and MODIS).

Figure 3.8 shows the spatial distribution of the FLUXNET observed annual ET and its differences with the model simulations. The spatial pattern of the FLUXNET annual ET generally follows the annual precipitation distribution, with very high ET in the Southeastern CONUS and low ET in the Western CONUS. One obvious exception is in the Pacific Northwest region where annual precipitation rate is very high while the annual

ET is not correspondingly high. Note that the snowfall in this region is much higher than in other regions, so the low temperature during snow seasons explains the low ET.
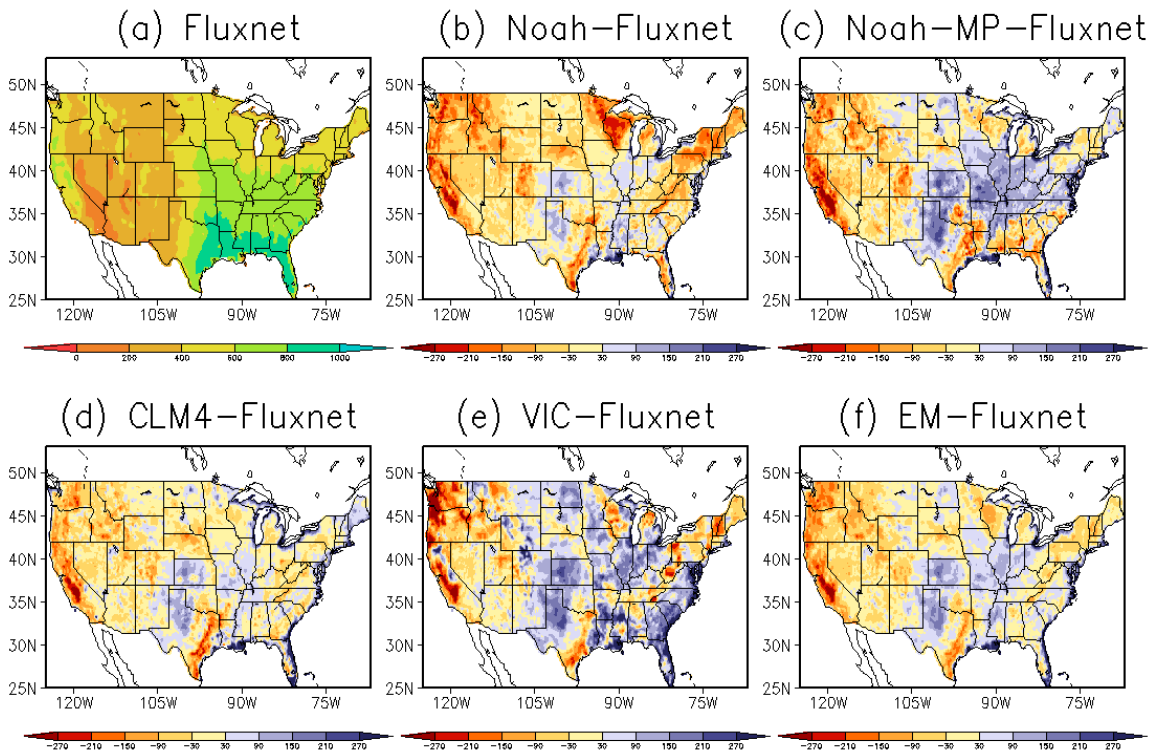
Figure 3.8 Spatial distribution of mean annual evapotranspiration (mm) from FLUXNET (a) and its differences with the model simulations for (b) Noah, (c) Noah-MP, (d) CLM4, (e) VIC, and (f) Ensemble mean (EM) of the four land models.

All four LSMs tend to underestimate ET for areas along the West Coast and in the Rocky Mountains while overestimating ET in the Texas Panhandle and Eastern Kansas. The Noah model underestimates ET for most of the CONUS, particularly in the West Coast, the Rocky Mountains, and the eastern Great Lakes regions. The Noah-MP model mainly underestimates ET for the Western U.S. and overestimates ET in the southern Midwest. The CLM4 model features small difference (shaded areas with light colors) with FLUXNET for most of the CONUS. The VIC model overestimates ET for a large portion

of the study domain, particularly in the Southeast region. Similar to CLM4, the model ensemble mean also shows small difference with the FLUXNET ET. Overall, CLM4 performs the best in ET simulation against FLUXNET observation while Noah-MP and VIC show intermediate improvement to Noah.
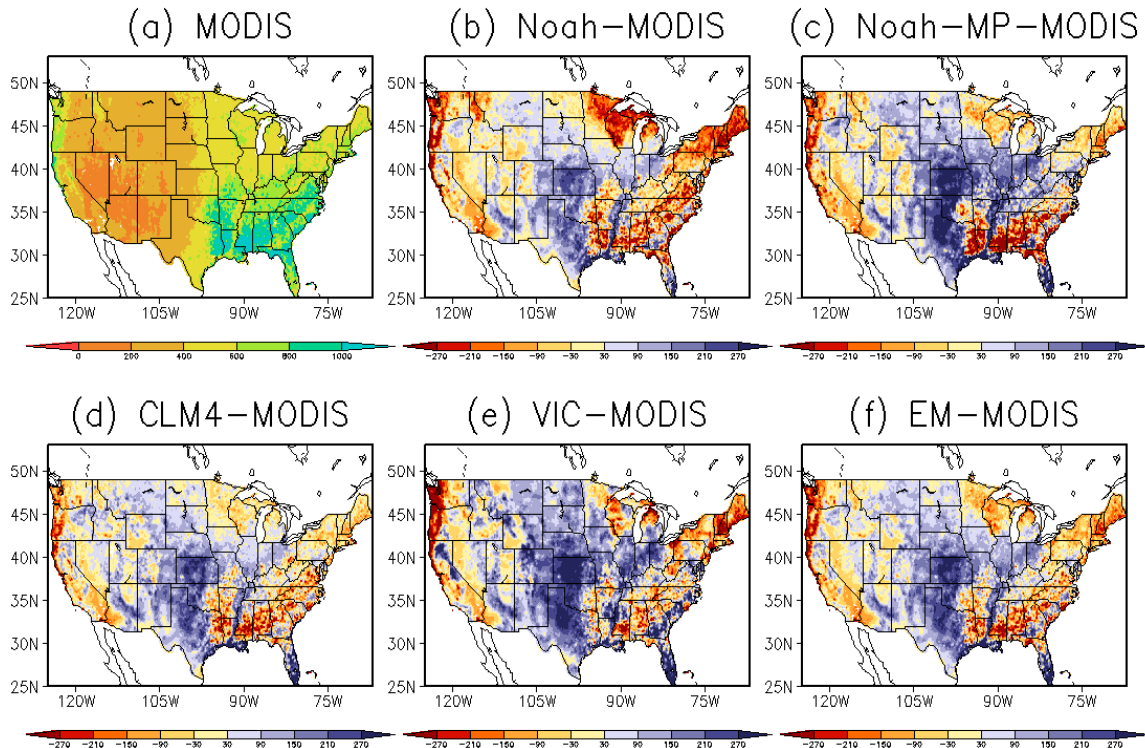
Figure 3.9 Same as Figure 3.8 except for replacing FLUXNET ET with MODIS ET.

Figure 3.9 is same as Figure 3.8 except for replacing FLUXNET ET with MODIS ET. The annual MODIS ET shows similar spatial patterns as FLUXNET ET. However, MODIS ET is generally less than FLUXNET ET for most areas, which results in larger overestimations by the four LSMs. One exception is in the area indicated by the red oval in Figure 3.9a, where the MODIS ET is very high (not seen in the FLUXNET ET). As a result, all LSMs produce lower ET than MODIS in this region.

3.5.4 Evaluation of Soil Moisture Simulation

Reliable gridded soil moisture observational data is not yet available, however, soil moisture measured at a site can reflect the temporal variation of soil water for its surrounding region, up to 500 km in radius (Entin et al., 2000). Figure 3.10 shows the comparison of monthly root-zone soil moisture (the top 1 m of the soil column) between SCAN observation and model simulations for the six regions indicated in Figure 3.1, with the statistics (correlation coefficient *r* and RMSE) of model performance summarized in Table 3.4. In the Southeast and Northeast regions with dense sites and good data availability from SCAN, all four LSMs perform well in terms of high correlation coefficient and low RMSE. The correlation coefficient values for Noah-MP, CLM4, and VIC are all above 0.72, which are at the higher end of the results reproduced by a range of LSMs as reported in the second Global Soil Wetness Project (Guo and Dirmeyer, 2006). This is slightly superior to the Noah model. Noah-MP (CLM4) produces the highest correlation coefficient for the Northeast (Southeast), while CLM4 (VIC) produces the lowest RMSE in the Southeast (Northeast) region. In the Texas region, the Noah-MP model stands out with the much higher correlation coefficient and lower RMSE compared to the other three models. In this region, the Noah-MP simulated soil moisture is very close to the SCAN observation while the other three LSMs simulated soil moisture is higher than the SCAN observation. In the remaining three regions, the data availability are low, therefore we can only use it as an approximate value without making some strong conclusions out of it. The model simulated soil moisture is smaller (larger) than the SCAN observation in the Great Plains region (Northwest and Southwest regions).
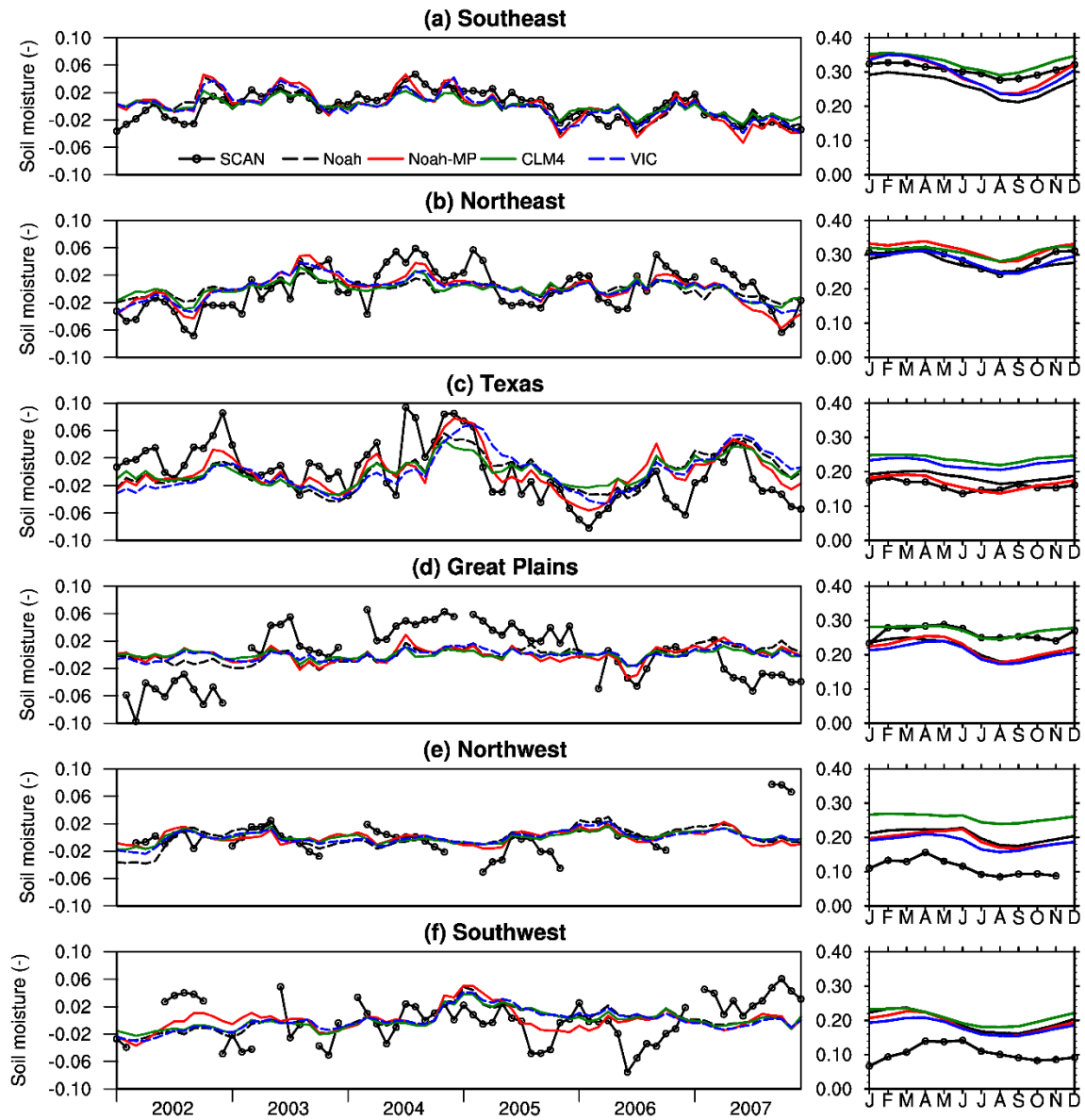
Figure 3.10 Monthly anomaly (left panels) and climatologically seasonal cycle (right panels) of spatially averaged root-zone soil moisture (the top 1 m of the soil column) from SCAN observation and model simulations for six regions: (a) Southeast, (b) Northeast, (c) Texas, (d) Great Plains, (e) Northwest, and (f) Southwest.

Table 3.4 Statistical summary of model performance in simulating soil moisture for the top 1-m soil, based on comparison with SCAN observation. Bold font indicates significant at the 95% confidence level. The thick underlines indicate the highest R^2 or lowest RMSE among the four land surface models.

Model	r							RMSE						
	SE	NE	TX	GP	NW	SW	Avg.	SE	NE	TX	GP	NW	SW	Avg.
Noah	**0.712**	**0.695**	**0.508**	**0.300**	0.120	0.070	0.401	0.015	0.025	0.035	<u>0.036</u>	0.022	0.032	0.027
Noah-MP	**0.781**	**0.762**	**0.716**	0.220	0.068	<u>0.181</u>	<u>0.455</u>	0.014	<u>0.020</u>	<u>0.028</u>	0.037	0.021	<u>0.030</u>	0.025
CLM4	<u>**0.794**</u>	**0.722**	**0.527**	0.166	<u>0.205</u>	0.044	0.410	<u>0.013</u>	0.024	0.034	0.037	<u>0.020</u>	0.031	0.027
VIC	**0.766**	**0.753**	**0.431**	<u>0.325</u>	0.177	0.008	0.410	<u>0.013</u>	0.021	0.038	0.036	<u>0.020</u>	0.032	0.027
Mean	0.763	0.733	0.546	0.253	0.142	0.076	0.419	0.014	0.022	0.034	0.036	0.021	0.031	0.026

Overall, the Noah simulated soil moisture is among the driest for all six regions; nevertheless, it is comparable with the SCAN observation. For the Noah-MP model, in addition to producing the highest correlation coefficient and low RMSE, its simulated soil moisture is closest to the SCAN observation. CLM4 simulated soil moisture correlates well with the SCAN observation; however, it produces the highest soil moisture among the four LSMs, much higher than the observation. This results in the very high RMSE for Texas, Southwest, and Northwest regions. Furthermore, the most important issue with CLM4 is that the variability of its simulated soil moisture is too low compared to the SCAN observation and the other LSMs, which is an issue found in CLM version 3 (Oleson et al., 2008). Although the issue is partly alleviated in CLM4 by adjusting the parameters that control the water table position (Lawrence et al., 2011), its soil moisture variability is still weaker than the observation and the other LSMs. Based on Niu et al. (2011), however, this issue is attributed to the parameterization of the soil moisture factor for stomatal resistance (the β factor). By running Noah-MP with different β factors, they demonstrated that the CLM-type β factor produced smaller root zone soil moisture variability than did the Noah-type β factor. For the VIC model, although it produces slightly lower (higher) correlation

coefficient (RMSE) than Noah-MP, its simulated soil moisture is close to Noah-MP and agrees well with the SCAN observation.

3.6 CONCLUSIONS

This study used the NLDAS test bed to assess the hydrological performance of four LSMs: Noah, Noah-MP, CLM4, and VIC. All models were driven by the same NLDAS-2 atmospheric forcing and evaluated against the same observational data sets. Compared to Noah, the other three models show significant improvements in TWS and streamflow and moderate improvements in ET and soil moisture. Among these LSMs, Noah-MP shows the best performance in simulating soil moisture and is among the best in simulating TWS; CLM4 shows the best performance in simulating ET; and VIC shows the best performance in simulating streamflow.

Deficiencies are also found in these LSMs. The Noah-MP simulated ET grows too fast in the spring, which coincides with the fact that its modeled LAI peaks too soon. Although the CLM4 model produced the highest correlation coefficient for TWS anomaly for the entire CONUS, it produced either too high or too low amplitude of the annual TWS variation. In addition, CLM4 produced much weaker soil moisture variability than both SCAN observations and other LSMs. Finally, the VIC model constantly overestimated ET compared to observations from MODIS and FLUXNET.

One must bear in mind that the confidence of this study may be undermined by the possible uncertainties from various sources. First of all, although the NLDAS atmospheric forcing is considered high quality, the error may not be negligible for some regions. Second, observational data have errors. For example, the water balance may not be closed for the NLDAS precipitation, USGS streamflow, and MODIS (or FLUXNET) ET, as they are measured separately and each one involves errors (Gao et al., 2010). This is part of the

reason why one model that performs well in simulating streamflow may not perform well in simulating ET at annual scales. Third, each model involves many physical processes and their interactions are complex, and thus, it is difficult to identify one effect from another.

In summary, this study gave an overview of the performances of four LSMs over CONUS. We demonstrated that, by providing reliable atmospheric forcing data, four in-house LSMs (more to be added), most appropriate observational data, and necessary tools, the NLDAS test bed is a valid platform for evaluating land models on continental or large river basin scales in the U.S.

3.7 ACKNOWLEDGEMENTS

Chapter 4: Integration of Nitrogen Dynamics into the Noah-MP Land Surface Model for Climate and Environmental Predictions

4.1 ABSTRACT

Climate and terrestrial biosphere models consider nitrogen an important factor in limiting plant carbon uptake, while operational environmental models view nitrogen as the leading pollutant causing eutrophication in water bodies. The community Noah land surface model with multi-parameterization options (Noah-MP) is unique in that it is the next generation land surface model for the Weather Research and Forecasting meteorological model and for the operational weather/climate models in the National Centers for Environmental Prediction. In this study, we add capability to Noah-MP to simulate nitrogen dynamics by coupling the Fixation and Uptake of Nitrogen (FUN) plant model and the Soil and Water Assessment Tool (SWAT) soil nitrogen dynamics. This incorporates FUN's state-of-the-art concept of carbon cost theory and SWAT's strength in representing the impacts of agricultural management on the nitrogen cycle. Parameterizations for direct root and mycorrhizal-associated nitrogen uptake, leaf retranslocation, and symbiotic biological nitrogen fixation are employed from FUN, while parameterizations for nitrogen mineralization, nitrification, immobilization, volatilization, atmospheric deposition, and leaching are based on SWAT. The coupled model is then evaluated at the Kellogg Biological Station—a Long-term Ecological Research site within the U.S. Corn Belt. Results show that the model performs well in capturing the major nitrogen state/flux variables (e.g., soil nitrate and nitrate leaching). Furthermore, the addition of nitrogen dynamics improves the modeling of the carbon and water cycles (e.g., net primary productivity and evapotranspiration). The model improvement is expected to advance the capability of Noah-MP to simultaneously predict weather and water quality in fully coupled Earth system models.

4.2 INTRODUCTION

Over the past several decades, eutrophication—high concentrations of nutrients in freshwater bodies leading to severe oxygen depletion from the resultant algal blooms—has become a worldwide problem facing river, lake and coastal waters (Howarth et al., 2006; Conley et al., 2009). As one of the greatest threats to freshwater and coastal ecosystems, eutrophic conditions lower biotic diversity, lead to hypoxia and anoxia, increase the incidence and duration of harmful algal blooms, and change ecological food webs that reduce fish production (National Research Council, 2000; Diaz and Rosenberg, 2008). These eutrophic conditions are attributed to excessive fertilizer leaching in river basins (Boyer et al., 2006; Boesch et al., 2009). To complicate this further, climate variation and climate change also determine the variation of hypoxia extent (Donner and Scavia, 2007): higher temperatures may extend the thermal stratification period and deepen the thermocline, thereby resulting in the upwelling of nutrients from sediment and increasing the concentration of nutrients in the bottom layer of water in lakes (Komatsu et al., 2007). Further, higher precipitation produces more runoff and very likely more nutrients are delivered to the ocean as well (Donner and Scavia, 2007).

Nitrogen (N) is recognized as the leading nutrient causing eutrophication. Without human interference, N cycling is relatively slow, as most ecosystems are efficient at retaining this in-demand nutrient. N enters soil regularly either through atmospheric wet and dry deposition or through atmospheric N_2 fixation by microorganisms (occurring mostly in legume plants). N taken up by plants is confined to relatively slow processes (e.g., growth, decay, and mineralization); in some regions or during the growing season, N may also limit plant growth, which reduces carbon sequestration over land (Fisher et al., 2012). In addition, N cycling produces nitrous oxide (N_2O) which is considered one of the important greenhouse gases responsible for climate warming. These facts make the N cycle

important for studying the response of the climate to the elevated greenhouse gas concentrations. With human tillage of soils, mineralization and nitrification of N is amplified, which results in the reduction of N storage in soil (Knops and Tilman, 2000; Scanlon et al., 2008). In addition, a large amount of N fertilizer is applied in specific areas within a short period of time; as a result, a massive excess of N is leached to the aquatic systems through discharge and erosion, which contributes to the eutrophication in aquatic systems.

Many of these N processes have been included in land surface, hydrologic, and water quality models developed particularly for environmental, climate, and agricultural applications (Dickinson et al., 2002; Thornton et al., 2007; Wang et al., 2007; Kronvang et al., 2009; Schoumans et al., 2009b; Bonan and Levis, 2010; Fisher et al., 2010). These developments are still in their infancy, and large scale climate models lack N leaching parameterizations that are comparable to those used in water quality models. Thus, large scale models are not feasible for inherently fine-scale applications such as agricultural fertilization management and water quality prediction. Therefore, the present study improves these weaknesses by incorporating the strength of agriculture-based models into large scale LSMs.

The community Noah LSM with multi-parameterization options (Noah-MP) (Niu et al., 2011; Yang et al., 2011) is used as an exemplar of LSMs because it is the next generation LSM for the Weather Research and Forecasting (WRF) meteorological model (Rasmussen et al., 2014) and for the operational weather and climate models in the NOAA/National Centers for Environmental Prediction. Because Noah-MP has an interactive vegetation canopy option, which predicts the leaf area index (LAI) as a function of light, temperature, and soil moisture, it is logical to augment this scheme with N limitation and realistic plant N uptake and fixation. The state-of-the-art vegetation N model

is the Fixation and Uptake of Nitrogen (FUN) model of Fisher et al. (2010), which is embedded into the Joint UK Land Environment Simulator (JULES) (Clark et al., 2011a) and the Community Land Model (CLM) (Shi et al., in preparation). Modeling the impacts of agricultural management (e.g., fertilizer use) on N leaching is the strength of the Soil and Water Assessment Tool (SWAT) (Neitsch et al., 2011). Therefore, this study incorporates into Noah-MP both FUN's strength in plant N uptake and SWAT's strength in soil N cycling and agricultural management.

Our objective is to develop and utilize a land surface modeling framework for simultaneous climate (carbon) and environmental (water quality) predictions. We first describe the nitrogen dynamic model which combines equations used in FUN and SWAT. We then focus on evaluating the new integrated model at a cropland site, because globally fertilizer application on croplands contributes approximately half of the total N input to soil, with the other half by natural processes (i.e. atmospheric deposition and biological N fixation) (Gruber and Galloway, 2008). Furthermore, cropland is a major source of N loading in water bodies. We evaluate the new model against observed soil moisture content, concentration of soil nitrate, concentration of nitrate leaching from soil bottom, and annual net primary productivity (NPP). We then analyze the impacts of the addition of N dynamics on the carbon and water cycles. To guide the use of this model on regional scale, we also analyze the impacts from different fertilizer application scenarios. Finally, we discuss other model behaviors, i.e., N uptake from different pathways and the major soil nitrate fluxes.

4.3 MODELS, DATA, AND METHODS

4.3.1 Noah-MP

The Noah-MP model was augmented from the original Noah LSM with improved physics and multi-parameterization options (Niu et al., 2011; Yang et al., 2011), based on

a state-of-the-art multiple-hypothesis framework (Clark et al., 2011b). Noah-MP provides users with multiple options for parameterization in leaf dynamics, canopy stomatal resistance, soil moisture factor for stomatal resistance, and runoff and groundwater. Until this work, Noah-MP did not include any N dynamics. The only N-related parameterization is in the calculation of the maximum rate of carboxylation (V_{max}, Equation 4.1)—an important factor in estimating the total carbon assimilation (or photosynthesis) rate (Niu et al., 2011).

$$V_{\max} = V_{\max 25} a_{v\max}^{\frac{T_v - 25}{10}} f(N) f(T_v)\beta \tag{4.1}$$

where $V_{\max 25}$ a is maximum carboxylation rate at 25°C (μmol CO_2 m^{-2} s^{-1}), $a_{v\max}$ is a temperature sensitive parameter, $f(T_v)$ is a function that mimics the thermal breakdown of metabolic processes, $f(N)$ is a foliage nitrogen factor ($f(N) \leq 1$), and β is the soil moisture controlling factor. Since there were no N dynamics in the model, $f(N)$ was set as a constant 0.67, which translates to a constant 33% of V_{max} down-regulation due to N stress. This factor was originally used in Running and Coughlan (1988) and first adapted into LSMs by Bonan (1991).

Our modifications to the original Noah-MP mainly concern the sub-models dealing with dynamic leaf and subsurface runoff. The dynamic leaf option is turned on to provide net primary productivity (NPP) and biomass to the newly coupled N dynamic sub-model. In the original Noah-MP model, subsurface runoff from each soil layer was not an explicit output, but it is now a new output in the updated model. However, N concentrations are different among soil layers, which affects the amount of N removed from each soil layer by subsurface runoff. Therefore, in conjunction with the runoff scheme options 1 (TOPMODEL with groundwater) and 2 (TOPMODEL with an equilibrium water table),

the lumped subsurface runoff for all four layers is first calculated, and then the water is removed from each soil layer weighted by hydraulic conductivity and soil layer thickness.

4.3.2 Nitrogen Dynamics

In Noah-MP, the soil N model structure is the same as in SWAT, which includes five N pools consisting of two inorganic forms (NH_4^+ and NO_3^-) and three organic forms (active, stable, and fresh pools). The N processes employed from SWAT are mineralization, decomposition, immobilization, nitrification, denitrification, and atmospheric deposition. The N processes employed from FUN are uptake and symbiotic biological N fixation, which can be further divided into active and passive soil N uptake, leaf N retranslocation, and symbiotic biological N fixation. Figure 4.1 shows the flow chart of the nitrogen dynamic model. Table 4.1 shows the model input variables and parameters.

Nitrogen Uptake and Fixation

Plant N uptake and fixation follows the framework of Fisher et al. (2010), which determines N acquired by plants through (1) advection (passive uptake), (2) retranslocation (resorption), (3) active uptake, and (4) symbiotic biological N fixation.

Noah-MP calculates the NPP or, following FUN, its available carbon, C_{NPP} (kg C m^{-2}). To maintain the prescribed carbon to nitrogen (C:N) ratio ($r_{C:N}$), the N demand, N_{demand} (kg N m^{-2}), is calculated:

$$N_{demand} = \frac{C_{NPP}}{r_{C:N}} \tag{4.2}$$

where the C:N ratio for the whole plant is computed from the for each component (leaf, root, and wood) of the plant proportionally to the biomass. C:N ratios for each component of the plant for each vegetation type are from Oleson et al. (2013).
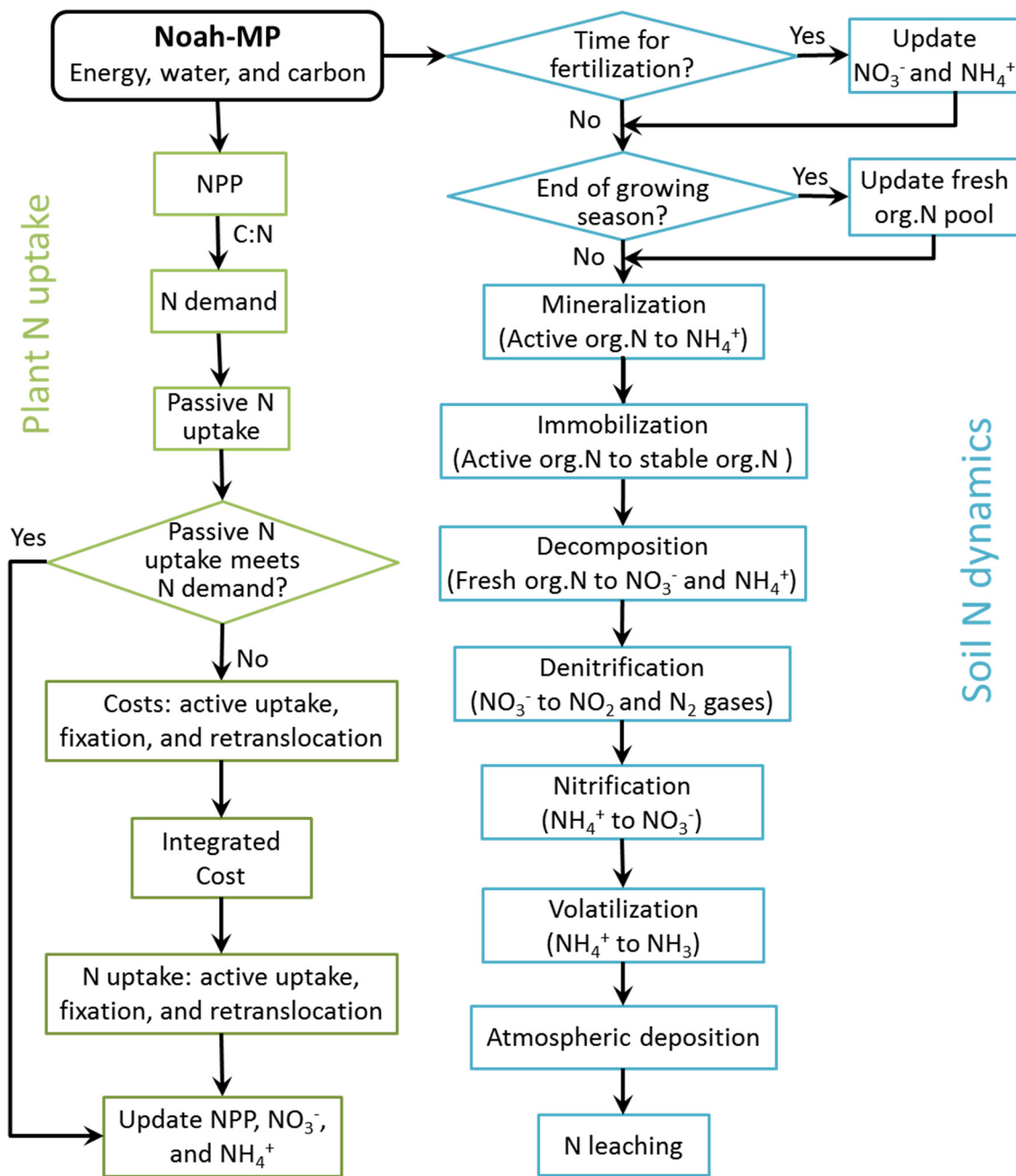
Figure 4.1 Flow chart of the nitrogen dynamic model.

Table 4.1 Model input variables and parameters.

Parameter	Definition	Controlling process	Unit	Value
$r_{C:N}$	C:N ratios for each component of the plant	Demand	–	Leaf 27, root 45
a	Empirical curve-fitting parameter	Fixation	–	–3.62
b	Empirical curve-fitting parameter	Fixation	–	0.27
c	Empirical curve-fitting parameter	Fixation	–	25.15
s	Scaling factor	Fixation	–	–62.5
k_N	Empirical curve-fitting parameter	Active uptake	kg C m^{-2}	1.0
k_C	Empirical curve-fitting parameter	Active uptake	kg C m^{-2}	1.0
N_{no3}	Initial value for NO$_3$ concentration in soil layer	Initialization	g N m^{-2}	6.7
N_{aon}	Initial value for humic organic N in soil layer	Initialization	g N m^{-2}	12.4
N_{fon}	Initial value for fresh organic N in soil layer	Initialization	g N m^{-2}	5.3
C_{org}	Initial organic carbon content in soil layer	Initialization	%	2.61, 0.35, 0.11, 0.07
ρ_b	Bulk density of the soil layer	Initialization	Mg m^{-3}	1.3, 1.4, 1.5, 1.6
f_{nh4n}	Fraction of mineral N in fertilizer that is ammonium	Fertilization	–	0.4
f_{surfn}	Fraction of fertilizer that is applied to the top 10 mm of soil	Fertilization	–	0.2
e_{mix}	Mixing efficiency of tillage operation	Tillage	–	0.3
β_{min}	Rate coefficient for mineralization of the humic organic nitrogen	Mineralization	–	0.002
β_{rsd}	Rate coefficient for mineralization of the fresh organic nitrogen in residue	Mineralization	–	0.04
β_{denit}	Rate coefficient for denitrification	Denitrification	–	1.4
$\gamma_{sw,thr}$	Threshold value of soil water factor for denitrification to occur	Denitrification	–	0.85
R_{no3}	Concentration of nitrate in the rain	Deposition	mg kg^{-1}	1.5
R_{nh4}	Concentration of ammonium in the rain	Deposition	mg kg^{-1}	1.0
D_{no3}	Constant of nitrate rate with dry deposition	Deposition	g N m^{-2} a^{-1}	0.2
D_{nh4}	Constant of ammonium rate with dry deposition	Deposition	g N m^{-2} a^{-1}	0.2
θ_e	Fraction of porosity from which anions are excluded	Leaching	–	0.15
β_{no3}	Nitrate percolation coefficient	Leaching	–	0.3

Note: Some parameters are not described in the paper. The values for C_{org} and ρ_b are for the four soil layers.

Because no extra energetic cost is needed, passive uptake, $N_{passive}$ (kg N m^{-2}), is the first and preferred source of N that a plant depletes.

$$N_{passive} = N_{soil} \frac{E_T}{s_d} \tag{4.3}$$

where N_{soil} is the available soil N for the given soil layer (kg N m^{-2}), E_T is transpiration rate (m s^{-1}), and s_d is the soil water depth (m). This pathway is typically a minor contributor except under very high soil N conditions.

If $N_{passive}$ is less than N_{demand}, then the remaining required N must be obtained from either retranslocation (N_{resorb}, kg N m^{-2}), active uptake (N_{active}, kg N m^{-2}), or BNF (N_{fix}, kg N m^{-2}), all of which are associated with energetic cost and hence require C expenditure (C cost). The C cost of fixation ($Cost_{fix}$, kg C kg N^{-1}), active uptake ($Cost_{active}$, kg C kg N^{-1}), and resorption ($Cost_{resorb}$, kg C kg N^{-1}) is calculated:

$$Cost_{fix} = s\{\exp[a + b \cdot T_{soil} \cdot (1 - 0.5 \cdot T_{soil}/c)] - 2\} \tag{4.4}$$

$$Cost_{active} = \left(\frac{k_N}{N_{soil}}\right)\left(\frac{k_C}{C_{root}}\right) \tag{4.5}$$

$$Cost_{resorb} = \frac{k_R}{N_{leaf}} \tag{4.6}$$

where a, b, and c (–3.62, 0.27 and 25.15, respectively) are empirical curve-fitting parameters (dimensionless) from Houlton et al. (2008), s is a scaling factor (=–62.5, use kg C kg N^{-1} °C for unit consistency), T_{soil} is soil temperature (°C), k_N and k_C are both 1 kg C m^{-2}, k_R is 0.01 kg C m^{-2}, and C_{root} is total root biomass (kg C m^{-2}). Active uptake is typically a dominant form of N uptake in natural ecosystems, consuming large quantities of NPP (that would otherwise go to growth or other allocations) in exchange for N.

Similar to parallel circuits, each carbon cost is treated as a resistor and the integrated cost ($Cost_{acq}$, kg C kg N^{-1}) is calculated (Brzostek et al., 2014):

$$\frac{1}{Cost_{acq}} = \frac{1}{Cost_{fix}} + \frac{1}{Cost_{resorb}} + \sum_{i=1}^{N_{soil}} \frac{1}{Cost_{active,i}} \tag{4.7}$$

where $Cost_{active,i}$ is the C cost for active N uptake of ith soil layer and N_{soil} is the total number of soil layers.

Using Ohm's law, N acquired from C expenditure (N_{acq}, kg N m^{-2}) is analogous to current and thus is calculated:

$$N_{acq} = \frac{C_{acq}}{Cost_{acq}}$$

(4.8)

Therefore, plant N uptake and fixation are computed and are updated for each N pool. In addition, the effect of N-limitation on CO_2 sequestration is represented in the model through the theory of C cost economics.

Mineralization, Decomposition, and Immobilization

Fresh organic residue is firstly broken down into simpler organic components via decomposition. The plant-unavailable organic N is then converted into plant-available inorganic N via mineralization by microbes. Plant-available inorganic N can also be converted into plant-unavailable organic N via immobilization by microbes.

Immobilization is incorporated into mineralization calculation (net mineralization). Mineralization and decomposition, which are only allowed to occur when temperature is above 0°C, are constrained by water availability and temperature. The nutrient cycling temperature factor for soil layer ly, $\gamma_{tmp,ly}$, is calculated:

$$\gamma_{tmp,ly} = 0.9 \cdot \frac{T_{soil,ly}}{T_{soil,ly} + \exp\left[9.93 - 0.312 \cdot T_{soil,ly}\right]} + 0.1$$

(4.9)

where $T_{soil,ly}$ is the temperature of soil layer ly (°C).

The nutrient cycling water factor for soil layer ly, $\gamma_{sw,ly}$, is calculated:

$$\gamma_{sw,ly} = \frac{\theta_{ly}}{\theta_{s,ly}}$$

(4.10)

where θ_{ly} is the water content of soil layer ly on a given day (mm H$_2$O), and $\theta_{s,ly}$ is the water content of soil layer ly at field capacity (mm H$_2$O).

The mineralized N from the humus active organic N pool, $N_{mina,ly}$ (kg N m^{-2}), is calculated:

$$N_{mina,ly} = \beta_{mina,ly} \left(\gamma_{tmp,ly} \cdot \gamma_{sw,ly} \right)^{1/2} \cdot N_{aon,ly}$$

(4.11)

where β_{mina} is the rate coefficient for mineralization of the humus active organic nutrients, and $N_{aon,ly}$ is the amount of N in the active organic pool (kg N m^{-2}).

The mineralized N from the residue fresh organic N pool, $N_{minf,ly}$ (kg N m^{-2}), is calculated:

$$N_{minf,ly} = 0.8 \cdot \delta_{ntr,ly} \cdot N_{fon,ly}$$

(4.12)

where $\delta_{ntr,ly}$ is the residue decay rate constant, and $N_{fon,ly}$ is the amount of N in the fresh organic pool (kg N m^{-2}).

The decomposed N from the residue fresh organic N pool, $N_{dec,ly}$ (kg N m^{-2}), is calculated:

$$N_{dec,ly} = 0.2 \cdot \delta_{ntr,ly} \cdot N_{fon,ly}$$

(4.13)

Nitrification and Ammonia Volatilization

Using a first-order kinetic rate equation, the total amount of ammonium lost to nitrification and volatilization in layer ly, $N_{nit|vol,ly}$ (kg N m^{-2}), is calculated:

$$N_{nit|vol,ly} = NH4_{ly} \cdot [1 - \exp(-\eta_{nit,ly} - \eta_{vol,ly})]$$

(4.14)

where $NH4_{ly}$ is the amount of ammonium in layer ly (kg N m^{-2}), $\eta_{nit,ly}$ is the nitrification regulator , and $\eta_{vol,ly}$ is the volatilization regulator. The calculation of $\eta_{nit,ly}$ and $\eta_{vol,ly}$ is described in Neitsch et al. (2011).

$N_{nit|vol,ly}$ is then partitioned to nitrification and volatilization. The amounts of N converted from NH_4^+ and NO_3^- of the ammonium pool via nitrification and volatilization are then calculated:

$$N_{nit,ly} = \frac{fr_{nit,ly}}{\left(fr_{nit,ly} + fr_{vol,ly}\right)} \cdot N_{nit|vol,ly} \tag{4.15}$$

$$N_{vol,ly} = \frac{fr_{vol,ly}}{\left(fr_{nit,ly} + fr_{vol,ly}\right)} \cdot N_{nit|vol,ly} \tag{4.16}$$

where $fr_{nit,ly}$ and $fr_{nit,ly}$ are the estimated fractions of N lost through nitrification and volatilization, respectively. They are calculated from the individual regulator in Equation 4.14 as following:

$$fr_{nit,ly} = 1 - \exp[-\eta_{nit,ly}] \tag{4.17}$$

$$fr_{vol,ly} = 1 - \exp[-\eta_{vol,ly}] \tag{4.18}$$

Denitrification

Denitrification is the process that bacteria remove N from soil (converting NO_3^- to N_2 or N_2O gases). Denitrification rate, $N_{denit,ly}$ (kg N m^{-2}), is calculated:

$$\begin{cases} N_{denit,ly} = NO3_{ly} \cdot [1 - \exp(-\beta_{denit} \cdot \gamma_{tmp,ly} \cdot orgC_{ly})] & \text{if } \gamma_{sw,ly} \geq \gamma_{sw,thr} \\ N_{denit,ly} = 0 & \text{if } \gamma_{sw,ly} \geq \gamma_{sw,thr} \end{cases} \tag{4.19}$$

where $orgC_{ly}$ is the amount of organic C in the layer (%), β_{denit} is the rate coefficient for denitrification, and $\gamma_{tmp,ly}$ is the nutrient cycling temperature factor for soil layer ly, $\gamma_{sw,ly}$ is the nutrient cycling water factor for soil layer ly, and $\gamma_{sw,thr}$ is the threshold value of $\gamma_{sw,ly}$ for denitrification to occur.

Atmospheric Deposition

While the mechanism of atmospheric deposition is not fully understood, the uncertainty is parameterized into the concentration of nitrate/ammonium in the rain for wet deposition, and the nitrate/ammonium dry deposition rate for dry deposition.

The amount of nitrate and ammonium added to the soil through wet deposition, $NO_{3,wet}$ (kg N m^{-2}) and $NH_{4,wet}$ (kg N m^{-2}), are calculated:

$$NO_{3,wet} = 0.01 \cdot R_{NO3} \cdot P$$

(4.20)

$$NH_{4,wet} = 0.01 \cdot R_{NH4} \cdot P$$

(4.21)

where R_{NO3} is the concentration of nitrate in the rain (mg N L^{-1}), R_{NH4} is the concentration of ammonium in the rain (mg N L^{-1}), and P is the amount of precipitation on a given day. Both R_{NO3} and R_{NH4} are constants in the model and their values used in this study are listed in Table 4.1.

Fertilizer Application

The N fertilizer application process is included in the new model as well. If real fertilizer application data (timing and amount for a specific year) are available, they can be used as model inputs. Otherwise, a fixed amount of N fertilizer (e.g., 7.8 g N m^{-2} a^{-1}) is applied at a fixed time of a year (e.g., June 20).

Leaching

N leaching from land to water bodies is a consequence of soil weathering and erosion processes. In particular, organic N attached to soil particles is transported to surface water through soil erosion. Therefore, the modified universal soil loss equation (USLE) (Williams, 1995) is used to determine soil erosion. The detail of the calculation is described in Neitsch et al. (2011).

N in nitrate form can be transported with surface runoff, lateral runoff, or percolation, which is calculated:

$$NO_{3,surf} = \beta_{NO3} \cdot conc_{NO3,mobile} \cdot Q_{surf}$$

(4.22)

$$\begin{cases} NO_{3,lat,ly} = \beta_{NO3} \cdot conc_{NO3,mobile} \cdot Q_{lat,ly} & \text{for toplayer} \\ NO_{3,lat,ly} = conc_{NO3,mobile} \cdot Q_{lat,ly} & \text{for lower layers} \end{cases}$$

(4.23)

$$NO_{3,perc} = conc_{NO3,mobile} \cdot w_{perc,ly}$$

(4.24)

where $NO_{3,surf}$, $NO_{3,lat,ly}$ and $NO_{3,perc}$ are the soil nitrates removed in surface runoff, in subsurface flow, and by percolation, respectively, (kg N m^{-2}), β_{NO3} is the nitrate percolation coefficient, $conc_{NO3,mobile}$ is the concentration of nitrate in the mobile water in for the layer (kg N mm H$_2$O^{-1}), and $w_{perc,ly}$ is the amount of water percolating to the underlying soil layer on a given day (mm H$_2$O), Q_{surf} is the surface runoff generated on a given day (mm H$_2$O), and $Q_{lat,ly}$ is the water discharged from the layer by lateral flow (mm H$_2$O).

4.3.3 Description of Evaluation Data and Model Configuration

At the regional scale, N related measurements are very limited. Even at site level, measurements are limited with respect to plant and carbon dynamics. The Kellogg Biological Station (KBS)—a Long-term Ecological Research (LTER) site—is unique in its long term continuous measurements of N related variables (soil nitrate, N leaching, mineralization, nitrification, and fertilizer application) in an agricultural setting with multiple crop and soil controls. Even within the LTER network, we cannot find a second site that conducts this kind of measurements. Therefore, the new model is evaluated at this site.

KBS is located in Hickory Corners, Michigan, USA, within the northeastern portion of the U.S. Corn Belt (42.40ºN, 85.40ºW, elevation 288 m). Mean annual temperature is

98

10.1ºC and mean annual precipitation is 1,005 mm with about half falling as snow. This study uses two treatments from this site: T1 cropland with conventional tillage and T2 cropland without tillage. Both treatments are rainfed are planted with the same crops: corn, soybean, and winter wheat in rotation.

This site features multiple N related measurements. Soil inorganic N concentration, which is sampled from surface to 25 cm soil depth, is available from 1989 to 2012. Concentration of inorganic N leaching at bedrock, which is sampled at 1.2 m of soil depth, is available from 1995 to 2013. These two measurements are used to evaluate model simulated concentrations of soil nitrate for the top 25 cm and nitrate leaching from the soil bottom. Soil N mineralization, which measures the net mineralization potential and is available from 1989 to 2012, is compared with the modeled mineralization rate qualitatively.

In addition, root zone soil moisture content is sampled and calculated on a dry weight basis. In order to compare with model output, it is converted to volumetric soil moisture by applying the soil bulk density. Annual NPP is converted from annual crop yields (1989–2013) by assuming a harvest index and a root to whole plant ratio for each crop type. The harvest indices for corn, soybean, and winter wheat are 0.53, 0.42, and 0.39, respectively. The root to shoot ratios for corn, soybean, and winter wheat are 0.18, 0.15, and 0.20, respectively (Prince et al., 2001; West et al., 2010). Although N uptake cannot be evaluated directly at this site, by evaluating the annual NPP, we can see the model's performance in representing the N limitation effect on plant growth.

Noah-MP requires the following atmospheric forcing data at least at a 3-hourly time step: precipitation, air temperature, specific humidity, surface air pressure, wind speed, incoming solar radiation, and incoming longwave radiation. The weather station at the site measures all of these except for incoming longwave radiation, but does not cover the entire

period from 1989 to 2014 (e.g., hourly precipitation data are only available since 2007), from when the N data are available. Therefore, atmospheric forcing data are extracted and used from the 0.125º × 0.125º gridded North American Land Data Assimilation System (NLDAS, (Xia et al., 2012c)) forcing data. Table 4.2 compares the atmospheric forcing data between NLDAS and site measurements for 2008–2014. We can see that the differences in precipitation and air temperature—the two most important forcing fields for N cycling—are very small, with relative biases –1.4% and 4.2%, respectively.

Table 4.2 Comparison of atmospheric forcing data between site observation and NLDAS (2008–2014).

Source	Precipitation (mm)	Air temperature (°C)	Relative humidity (%)	Pressure (hPa)	Shortwave radiation (W m^{-2})	Wind speed (m s^{-1})	Wind direction (°)
Site obs.	937.19	9.15	73.44	982.29	157.07	3.37	194.72
NLDAS	924.45	9.55	76.50	983.47	171.03	4.74	206.43

Finally, the site management log records the detailed operational practices such as soil preparation, planting, fertilizer application, pesticide application, and harvest. N fertilizer application data include the date of application, rate, fertilizer type, and equipment used. The fertilizer application date and rate are used as model inputs.

4.4 RESULTS AND ANALYSES

4.4.1 Soil Moisture Content

Modeled volumetric soil moisture, which is important for nutrient cycling and plant growth, is compared to measured soil moisture (Figure 4.2). The model performs particularly well on both treatments (i.e., with and without tillage) (see, for example, Cai et al. (2014b)). The model simulated multiple year averages are both 0.243 for the two

treatments. These are very close to observations which are 0.238 and 0.264 for T1 and T2, respectively. The correlation coefficient is 0.78 for T1 and 0.76 for T2, which are considered high skills, especially on a daily scale.
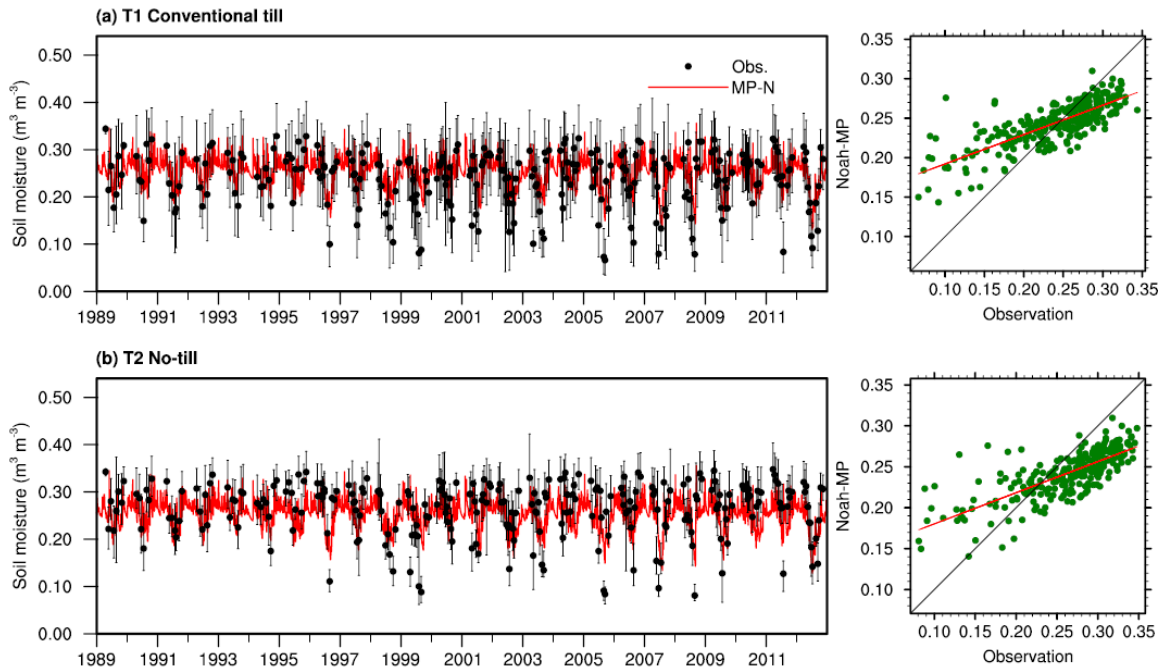
Figure 4.2 Observed and model simulated volumetric soil moisture from 1989 to 2012 for (a) Treatment 1: cropland with conventional tillage and (b) Treatment 2: cropland without tillage. The error bars show the observational ranges from up to six replicates for each treatment.

However, differences between modeled and observed soil moisture are also found. From observation (Figure 4.2), we can see that the treatment without tillage (T2) has slightly higher soil moisture than the treatment with tillage (T1). Therefore, tillage practice has impacts on soil moisture. However, the difference in modeled soil moisture is negligible between the two treatments (both are 0.243). This is because Noah-MP does not consider water redistribution due to tillage, although N redistribution is considered in the soil N dynamic sub-model. In addition, observed soil moisture has higher variations. As

we can see from Figure 4.2, observation tends to have either higher peaks or lower valleys than model simulation. We also notice that some values from observation are extremely low, which may not be necessarily true in reality. Considering the wide spread of the observational ranges defined by up to six replicating plots, Noah-MP provides a reasonable water environment for the N cycling.

4.4.2 Soil Nitrate

Soil nitrate concentration is the outcome of all N related processes that occur in soil such as decomposition, mineralization, nitrification, denitrification, and uptake. It determines the available N that plants can use. The skills in modeling the soil nitrate concentration reflect the overall performance of the model in simulating the N cycle. Figure 4.3 shows the comparison of the model simulated soil nitrate concentration with site observations for both T1 and T2. The model captures the major variations of the soil nitrate. N fertilizer application is responsible for the high peaks. These high peaks drop very fast at first and then drop slowly, which can sustain crop growth for few months.

The multi-year average of modeled soil nitrate concentration is 5.77 mg/kg (4.90 mg/kg) for T1 (T2), which is consistent with the observed 5.61 mg/kg (4.81 mg/kg). Correlation coefficients are 0.58 and 0.56 for T1 for T2, respectively. From the wide spread of the range error bars, we can see that soil N dynamics may be affected by a variety of complicated factors, which makes it difficult to model. Therefore, although the correlation coefficients are not considered high skills relative to the soil moisture statistics, they are still acceptable.

Figure 4.3 Observed and model simulated soil nitrate concentration from 1989 to 2011 for (a) Treatment 1: cropland with conventional tillage and (b) Treatment 2: cropland without tillage. The error bars show the observational ranges from up to six replicates for each treatment.

While both treatments show very similar patterns (Figure 4.3), T1 with conventional tillage tends to have higher soil nitrate concentration. This is understandable because tillage practices redistribute water and nutrients in soil, which accelerates the N cycling. Table 4.3 shows annual averages of major N fluxes for both treatments. T1 has higher rates of humus mineralization and residue decomposition, but at the same time, it also has higher rates of denitrification and leaching. Therefore, it produces more N_2O (a greenhouse gas) and more N runoff to rivers. Particularly, with higher N leaching, less soil nitrate is available for passive uptake by plant. As a result, plants need to acquire more N through active uptake.

Table 4.3 Annual averages of Noah-MP simulated major nitrogen fluxes and net primary productivity. The NPP within the parentheses is from observation.

| Treatment | NPP (gC/m²) | Uptake | | | | Humus Mineral. (gN/m²) | Residue Decomp. (gN/m²) | Denitrif. (gN/m²) | Leaching (gN/m²) |
		Passiv (gN/m²)	Active (gN/m²)	Fixation (gN/m²)	Retransl. (gN/m²)				
T1	432 (437)	6.18	0.90	2.88	0.50	3.79	12.30	10.48	7.19
T2	441 (471)	6.62	0.69	2.84	0.50	2.64	9.34	8.80	4.77

4.4.3 Nitrate Leaching from Soil Bottom

N leaching can be transported to rivers through surface and subsurface runoff and to groundwater through percolation from soil bottom. Only the last pathway is measured at this site. Figure 4.4 shows the comparison of concentrations of the leached solution from the soil bottom between model simulation and observation. The averaged concentration of N leaching from the soil bottom for T1 (T2) is 12.84 mg/kg (8.86 mg/kg) from model simulation and 13.57 mg/kg (9.26 mg/kg) from observation. The correlation coefficients are 0.43 and 0.40 for T1 and T2, respectively. Although these skills may not be considered satisfactory, the model can still produce comparable results with observation.

The peak in 2003 is extremely high and long lasting. This is probably due to the abnormal pattern of precipitation distribution in 2003. In a normal year, storms higher than 50 mm usually occur in either summer or fall. However, in 2003, there was an early storm on April 4 which reached 61 mm in one day. As we can see from Figure 4.4, the soil nitrate concentration is also high. The combination of high water infiltration (due to storm) and high soil nitrate concentration resulted in a large amount of soil nitrate being brought to the bottom soil layer. A few months following that, there was no large storm, which was again different from a normal year. As a result, the high concentration nitrate solution was drained slowly out of the bottom layer of soil. The modeled nitrate leaching also shows a peak over this period but the values are only close to the lower bound of the observed range.

This suggests that improvement is needed so the model can better capture peaks under this situation.
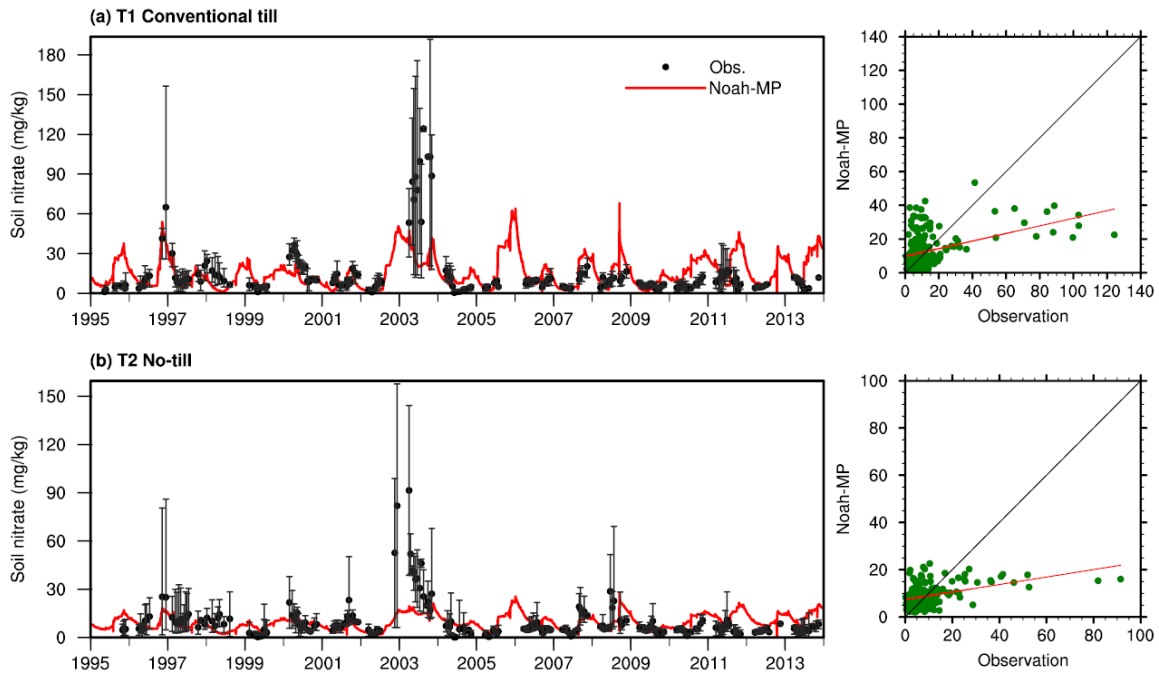
Figure 4.4 Observed and model simulated nitrate leaching from bottom of soil profile from 1995 to 2013 for (a) Treatment 1: cropland with conventional tillage and (b) Treatment 2: cropland without tillage. The error bars show the observational ranges from up to six replicates for each treatment.

Due to the low sampling frequency, observation tends to miss some high peaks of N concentration in leaching. One apparent example is the peak during late 2006 and early 2007, which can be captured by the model, but no observation is available during this period.

We also notice that without tillage, N leaching is about one third lower than that with tillage. Without tillage, the temporal variation is also smaller.

4.4.4 Annual NPP

NPP indicates the amount of C that is assimilated from the atmosphere into plants and thus is important in studying not only crop and ecosystem productivity, but also climate change feedbacks. NPP is mainly determined by plant photosynthesis and autotrophic respiration. It is also affected by water and nutrient stresses. In this study, N stress on plant growth is calculated by the reduction of NPP due to N acquisition, which can be considered another form of plant respiration. Figure 4.5 shows the comparison of simulated annual NPP against observation. Since the original Noah-MP without N dynamics also simulates NPP, its results are also shown here as a reference. The mean annual NPP simulated by the original Noah-MP is 544 gC/m^2 (the same simulation for both treatments as original Noah-MP does not distinguish tillage and no tillage). By including the N dynamics, simulated annual NPP is reduced to 432 gC/m^2 (441 gC/m^2) for T1 (T2), which is more consistent with observed 437 gC/m^2 (471 gC/m^2). The correlation coefficient increased to 0.77 for T1, and from 0.30 to 0.72 for T2, which is a significant improvement. This improvement is due to the better characterization of the amount of carbon allocated to N acquisition instead of growth.

Modeled rate of NPP down-regulation—the fraction of NPP reduction due to N limitation—is 35.4% and 34.7% for T1 and T2, respectively. These rates are close to the 33% of down-regulation rate used in the default Noah-MP. By dynamically simulating the demand and supply of N with time, these become even closer to the observations.

Surprisingly, even with slower N cycling, T2 produces slightly higher NPP (Table 4.3), which is consistent between model and observation. If this is the case, except for producing more N_2O gas and N runoff, is there any benefit from tillage? The answer is yes. Less N fertilizer is needed for cropland with tillage. Based on the site management log, in

total there was 194.8 gN/m^2 of N fertilizer applied to T1 from 1989 to 2013, which is less than the amount (210.7 gN/m^2) applied to T2 during the same period.
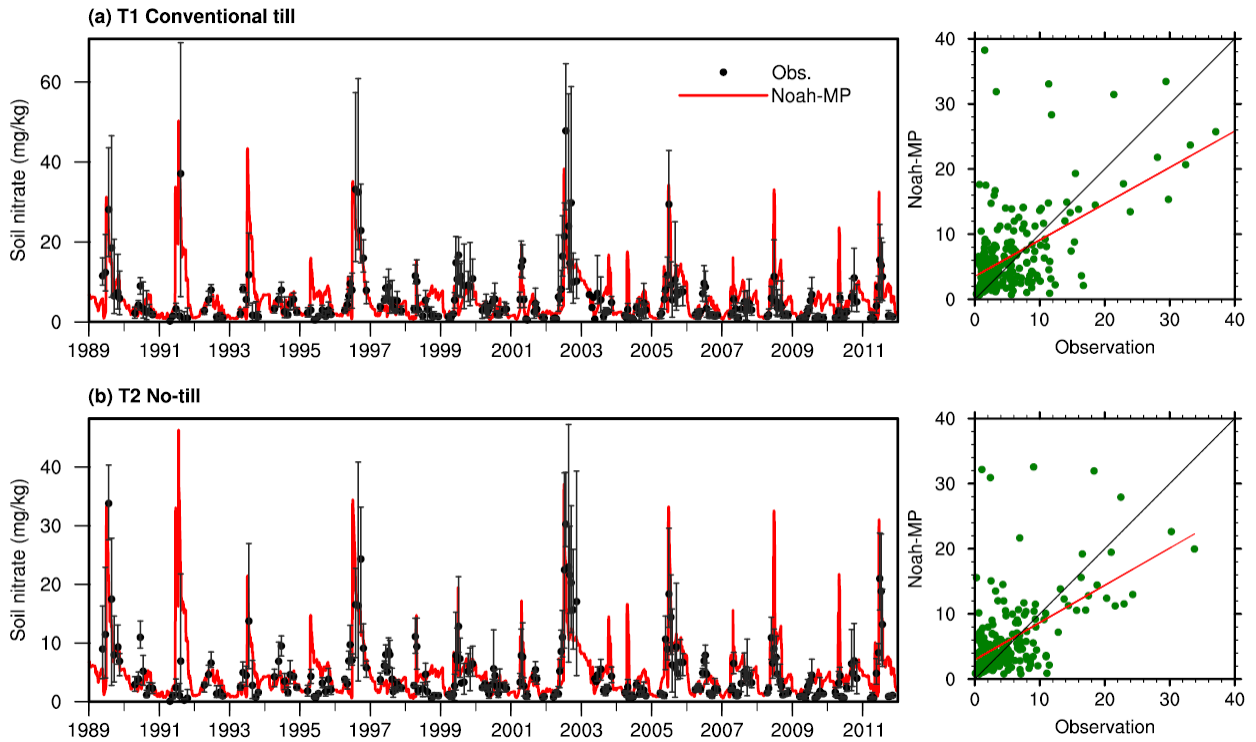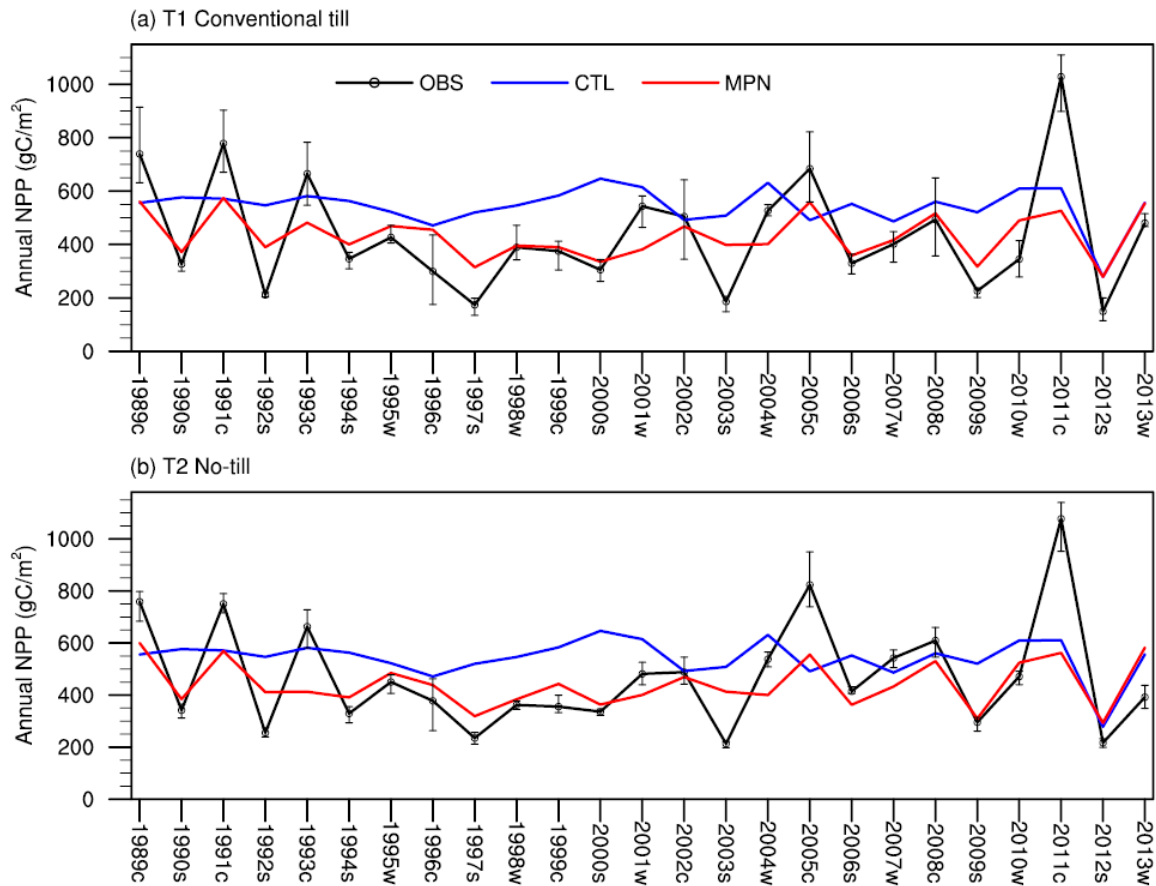

Figure 4.5 Observed and modeled annual NPP from 1989 to 2013 for (a) Treatment 1: cropland with conventional tillage and (b) Treatment 2: cropland without tillage. The error bars show the observational ranges from up to six replicates for each treatment. CTL: original Noah-MP without N cycling. MP-N: Noah-MP with N cycling. The letters in the x-axis labels are the crops harvested in that year (c: corn; s: soybean; w: winter wheat).

4.4.5 Impacts on Carbon Cycle

The coupling of the N dynamics into Noah-MP not only adds N related modeling but also affects other components of the model, i.e. the C and water cycles. This is because

the change in NPP affects leaf biomass and hence LAI. The change in LAI can affect photosynthesis, which in return affects NPP.

Figure 4.6 shows the comparison of the simulated C related state and flux variables between the default and N dynamics enhanced Noah-MP. We can see that NPP is decreased from 544 gC/m^2 to 432 gC/m^2. Most of the decrease occurs before the peak growing season, which results in a slight decrease in LAI. However, the peak LAI has very minor increase. After the peak, LAI decreases more slowly than the default, which is due to the decreased turnover rate proportional to the NPP down-regulation rate. If the turnover rate is not down-regulated accordingly, the peak LAI will be cut in half. Due to the slower turnover rate, more leaf biomass (indicated by LAI) is involved in photosynthesis. Therefore, compared to the default, Noah-MP with N dynamics generates higher gross primary production (GPP) during the second half of the growing season, although it is lower during the first half of the growing season. Annual mean GPP is decreased by about 28 gC/m^2.

Net ecosystem exchange (NEE) has a similar change. The annual NEE is –179 gC/m^2 (–183 gC/m^2) from Noah-MP with N dynamics (default Noah-MP), which is comparable to the NEE in West et al. (2010) for this region. Its absolute value is decreased by 4 gC/m^2, which means that C sink is slightly decreased. This decrease is small compared to the GPP decrease, probably because soil respiration is also decreased. All annual peaks of NPP, LAI, GPP, and NEE are delayed for about half month. This is probably due to the fact that the primary N fertilizations (> 10 gN/m^2) were mainly applied after late June and thus plants encountered high N stress during the first half of growing season.

Figure 4.6 (left column) Monthly and (right column) climatologically seasonal cycle of model simulated (a) LAI, (b) NPP, (c) GPP, and (d) NEE from default Noah-MP and enhanced Noah-MP with N dynamics. The values in the right column indicate annual mean for each term (black: default; red: N dynamics).

4.4.6 Impacts on Water Cycle

Through the changes in LAI and soil organic matters (SOM), the addition of N dynamics affects not only the C cycle but also the water cycle. The change in SOM is not currently considered and therefore, the impacts on the water cycle are from the change in LAI only, as shown in Figure 4.7. These impacts are most pronounced on plant transpiration, which is increased by 33 mm/a. The increase mostly occurs during and after the peak growing season. In Cai et al. (2014a), Noah-MP simulated evapotranspiration (ET) over croplands increases too fast during the first half of the growing season and reaches peak about one month earlier than observation. The delayed peaks of LAI and ET

can partly mitigate this issue. As there is more water extracted from soil by transpiration, soil moisture further decreases during the second half of growing season. Therefore, less water is available and thus soil evaporation is decreased by 9 mm/a. With the increase in ET, runoff is decreased by 13 mm/a.

Therefore, besides the great implications for C modeling and the potential for being used in environmental predictions, the addition of N dynamics can improve the hydrological simulations as well.
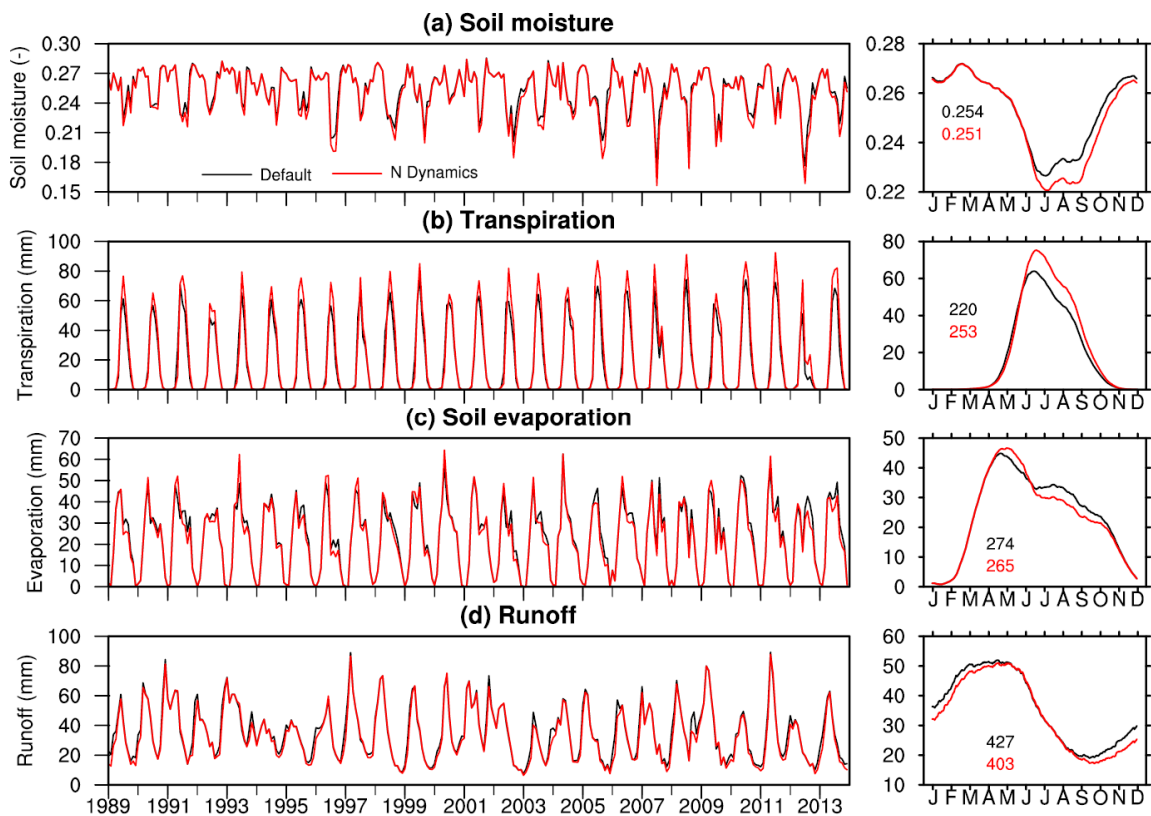
Figure 4.7 Same as Figure 4.6 except for (a) soil moisture, (b) transpiration, (c) soil evaporation, and (d) runoff.

110

4.4.7 Impacts of N Fertilizer Application

Observed N fertilizer application data is used in this study. However, this type of data is not always available, especially when models are applied in large regions. Often we only know the approximate amount of N fertilizer applied, without information on the exact dates. To guide the future large-scale application of this model, two additional experiments are run: (1) N fertilizer is applied on June 20 every year and (2) N fertilizer is applied on April 15 every year. The first experiment is designed because a large amount of N fertilizer is applied mostly during mid-June and early July. Other dates are also reported in the literature, therefore, we use April 15 as another example. Both experiments use the same amount of N fertilizer as the management log, which on average is 7.8 g N m^{-2} a^{-1}.

Figure 4.8 shows comparison of some of the most relevant results between the two experiments and the default with recorded dates and amount of N fertilizer application. Despite the different application time, the two experiments produce very consistent NPP with the default. The June 20 experiment is much closer to the default, even the seasonal variation is identical to the default. The largest discrepancy is in 1993 and 1996. Based on the management log, in these two years, a large amount of N fertilizer was applied, which resulted in much higher NPP than results from the two experiments. Since April 15 is much earlier than the primary fertilizer application dates, NPP from this experiment is flattened out through the year. This also confirms the statement in Section 4.4.5 that later N fertilizer applications delay plant growth. Simulated N uptake from both experiments shows exactly the same story as NPP.

The simulated N leaching shows the opposite pattern to NPP. The default simulation produces the highest leaching, followed by the June 20 experiment and then the April 15 experiment. This is very likely because the fertilizer application dates are closer to the flood season for the former two cases and the chance of fertilized N being flashed

111

out is higher. The difference in N fertilization dates also clearly affects the simulations of total soil nitrate. In the June 20 experiment, soil nitrate reaches the lowest level in May because no N fertilizer is applied before June 20. In the default case, N fertilizer can actually be applied as early as April, but with a smaller amount before mid-June, which prevents the soil nitrate concentration from getting too low. Besides a large amount of N fertilizer applied in later months, the other reason that the default simulation reaches the highest concentration of soil nitrate is because it produces higher NPP, which can be returned to soil for decomposition.
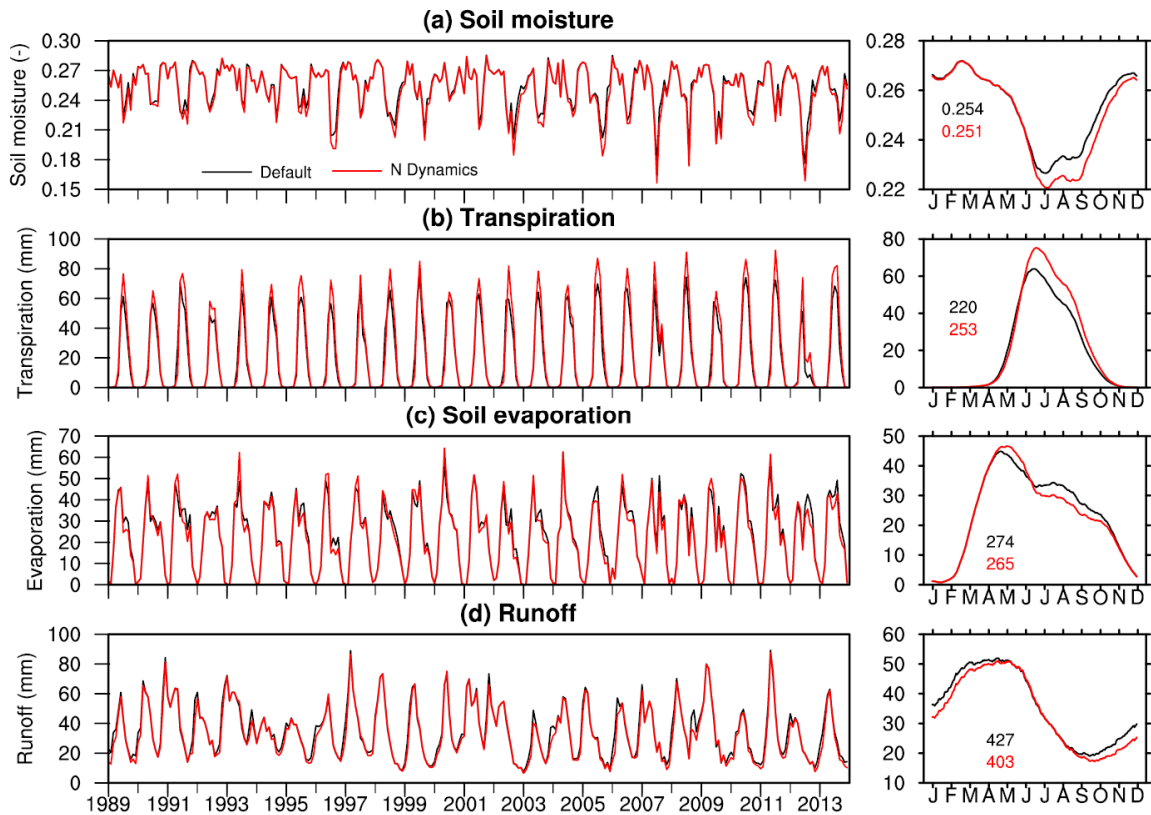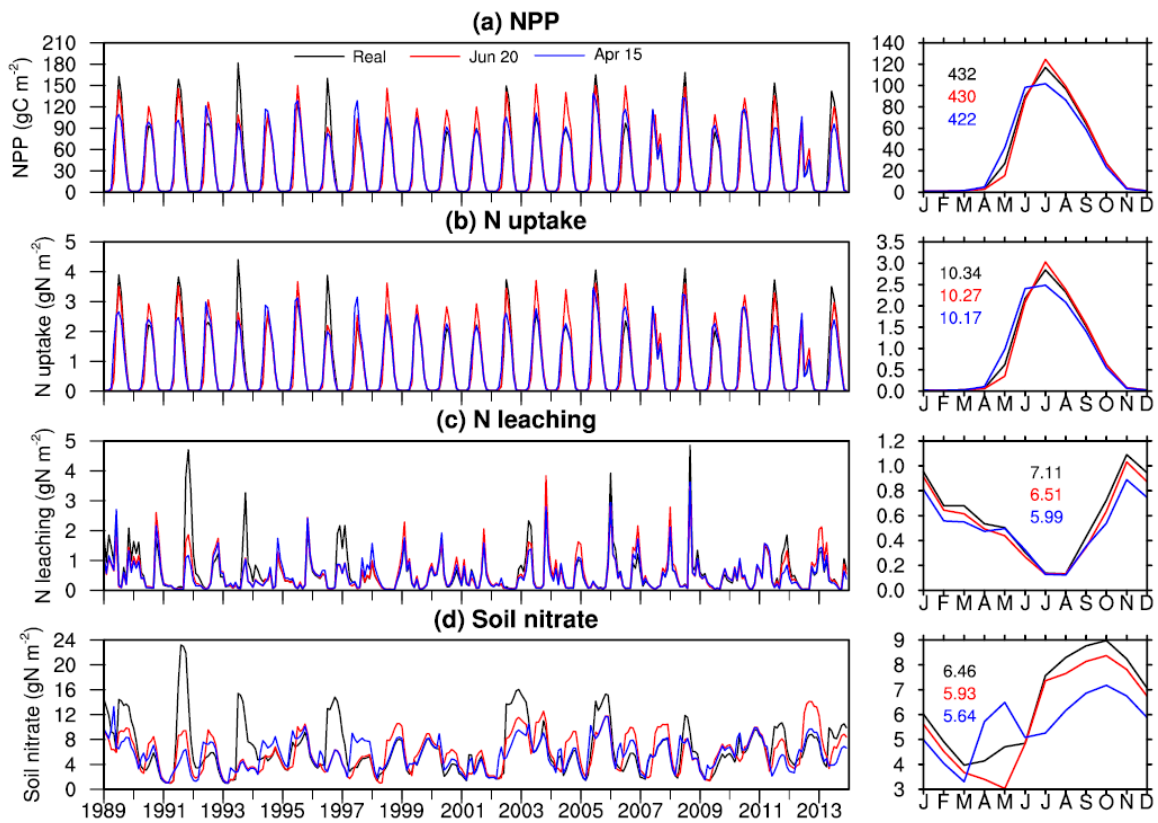

Figure 4.8 (left column) Monthly and (right column) climatologically seasonal cycle of model simulated (a) NPP, (b) N uptake, (c) N leaching, and (d) soil nitrate with different dates for N fertilization: real, June 20, and April 15. The values in the right column indicate annual mean for each term (black: real; red: June 20; blue: April 15).

112

Overall, the default simulation grows better plants (higher NPP) because N fertilizer is applied based on expert judgment of plants' demand. At the same time, however, it produces more N leaching. The experiment with closer dates of N fertilizer application to reality can better reproduce the N dynamics in observation. Therefore, although we cannot always know the exact dates of N fertilizer application, a survey on this can help to improve model simulation.

4.4.8 Analysis of N Uptake

As described in Section 4.3.2, plants can get N for growth from four pathways: passive uptake, active uptake, fixation, and retranslocation, and the last three require C costs. Figure 4.9 shows the actual N uptake from these pathways and their percentages of contribution to the total N uptake. Passive uptake is the dominant pathway, which contributes 57.7% of the total N uptake. Fixation, active uptake, and retranslocation contribute 28.6%, 8.7%, and 5.0%, respectively. This contrasts the results from the study by Brzostek et al. (2014) for non-fertilized trees, in which passive uptake only accounts for a small contribution. This is understandable because the purpose of fertilization is to minimize active uptake so that more NPP can be retained for crop growth. As demonstrated in Timlin et al. (2009), a higher fertilization rate results in a higher ratio of N uptake in transpiration to total N uptake. On the one hand, fertilization maintains soil nitrate concentration at high level. On the other hand, higher NPP for crop growth in turn results in higher LAI and thus higher transpiration. During peak growing season, therefore, plants receive a large amount of N under the combination of high transpiration and high soil nitrate concentration. During other periods, biological N fixation dominates.
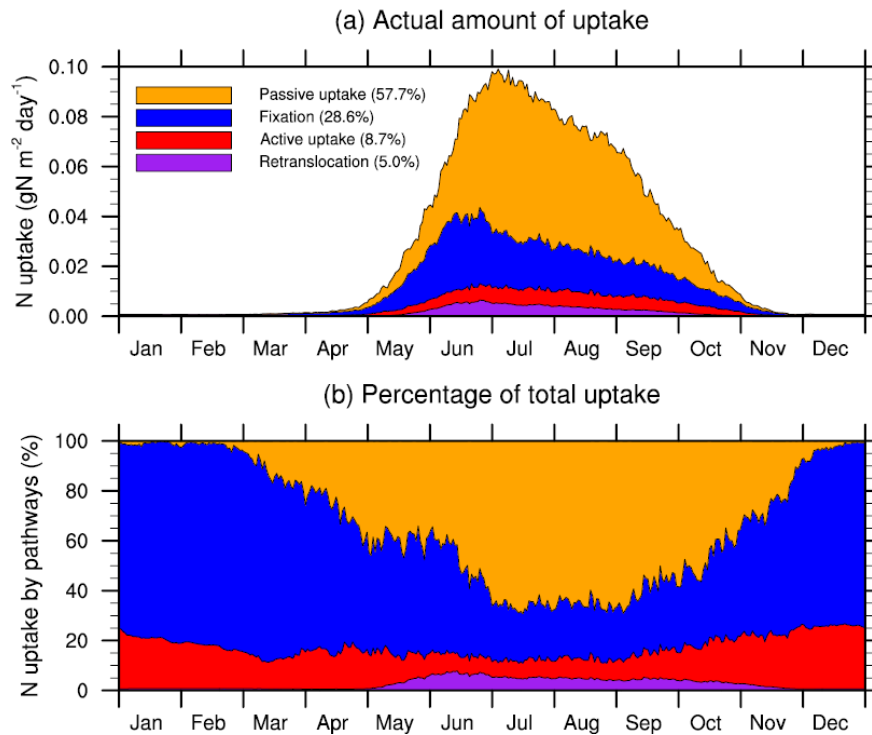
Figure 4.9 Daily climatology (1989–2013) of nitrogen uptake by pathways expresses as (a) actual amount of uptake and (b) percentage of total uptake.

4.4.9 Analysis of Major Soil Nitrate Fluxes

The soil nitrate storage with time is an outcome of the variations in incoming and outgoing fluxes. Besides N fertilizer and atmospheric deposition, humus mineralization and residue decomposition are the two major incoming fluxes. Because N fertilizer is a jumping behavior and atmospheric deposition is a relatively small fraction in this study ($\sim$1.5 $gN/m^2/a$), they are not analyzed here. The major outgoing fluxes are denitrification, leaching, and plant uptake.

Figure 4.10 shows the seasonal variation of the above major fluxes. During the growing season, N fertilizer provides an important role in meeting the plant N demand, however, residue decomposition still makes the largest contribution and is the dominant factor responsible for the increase in total soil nitrate. During the non-growing season, a

114

large amount of N is lost through denitrification and N leaching. However, when it reaches the peak growing season, plants consume a large fraction of soil nitrate, which leaves very little for denitrification and leaching. N leaching is mostly associated with the timing and intensity of precipitation. Denitrification is also associated with precipitation, but it is directly related to the soil water content. High denitrification rate occurs during high soil water content, especially during water logging.
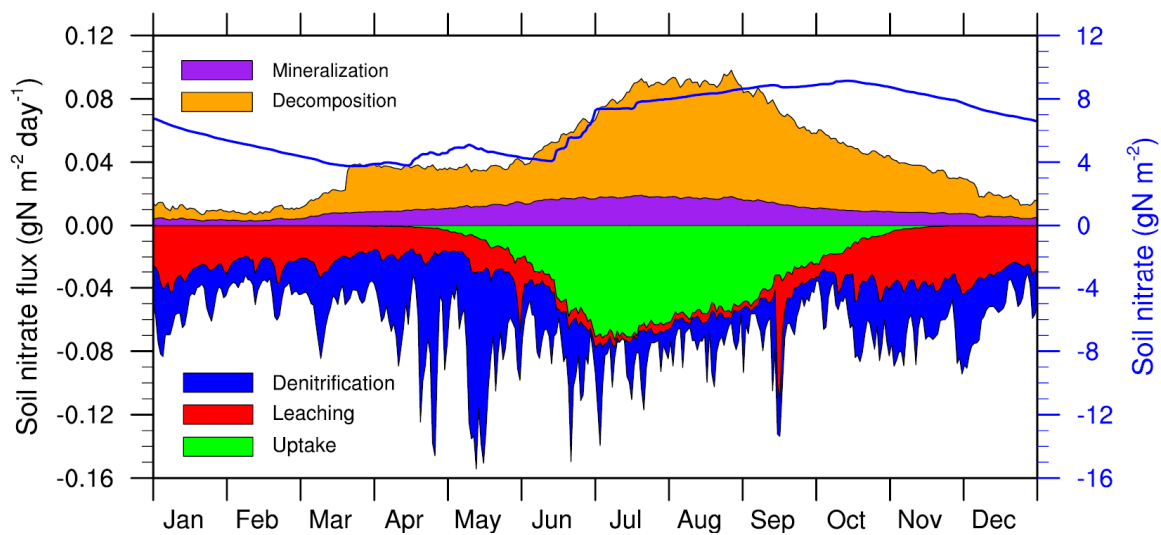
Figure 4.10 Daily climatology of the soil nitrate (blue solid line) and some major fluxes (color label bars) going in (humus mineralization and residue decomposition) and out (plant uptake, nitrate leaching, and denitrification) of the soil nitrate pool.

4.5 CONCLUSIONS

In this study, a dynamic N model is coupled into Noah-MP by incorporating FUN's strength in plant N uptake and SWAT's strength in soil N cycling and agricultural management.

We evaluated the new model at KBS that provides good-quality, long-term observed N and ecological data. The model simulated soil moisture is consistent with observation, which shows that Noah-MP provides a good water environment for the N

cycling. The simulated concentrations of soil nitrate and N leaching from soil bottom also compare well with observations. Although the model does not simulate some peaks well, especially for N leaching, the averages are very close to the observed values and the correlation coefficients are reasonable. Considering the wide spread of the range error bars defined by the measurements at the six replicates, the model shows high skill in capturing the major N flux/state variables. The significant improvement of annual NPP simulation demonstrates that the N limitation effect on plant growth is well represented in the model.

Moreover, the addition of N dynamics in Noah-MP improves the modeling of the C and water cycles. Comparing to the default Noah-MP, NPP simulations are improved significantly, in terms of consistent averages and much higher correlation coefficients with observation. The temporal pattern of simulated ET is also improved, featuring lower ET during spring and delayed peak.

This enhancement is expected to facilitate the simultaneous predictions of weather and environment by using a fully coupled system.

4.6 ACKNOWLEDGEMENTS

Chapter 5: Conclusions

5.1 CONCLUSIONS

This dissertation used Noah-MP as an example LSM to: (1) examine the effect of its recent development in improving hydrological simulations over the Mississippi River Basin; (2) investigate the advantages and disadvantages of Noah-MP in assessing the water balance over CONUS; (3) couple the FUN plant model and SWAT's soil nitrogen dynamics into Noah-MP and evaluate how the coupled model characterizes the major nitrogen fluxes and state variables; and (4) analyze how nitrogen dynamics affect the carbon and water simulations.

Chapter 2 reported that, with groundwater dynamics included in Noah-MP, WTD takes longer to reach equilibrium than without groundwater dynamics. This long WTD spin-up time influences the spin-up times of other variables because, when the water table is far from an equilibrium state, other variables such as runoff, ET, and soil moisture need to be adjusted to help WTD reach equilibrium. For the entire MRB, at least 34 years are required for the model to spin-up. For some mountainous regions with very deep water tables, hundreds of years may be required for the model to spin-up.

Runoff is found to be sensitive to three parameters: surface dryness factor (α), saturated hydraulic conductivity (k), and saturated soil moisture (θ_{max}); these three factors are selected for model calibration to improve runoff simulation. Although lumped calibration can improve model performance, distributed calibration is needed to obtain the best parameter values for some wet regions. If time and resources are limited for conducting automatic calibration (e.g., in this study), a better understanding of model physics and more analyses of observational data would shorten the calibration time and benefit the model performance.

117

Noah-MP has shown significant improvements in hydrological modeling. (1) The Noah-MP simulated runoff is significantly improved compared with the baseline Noah LSM output in the NLDAS-2 framework. The spatial pattern of the Noah-MP simulated runoff matches fairly well with both the UNH-GRDC runoff and the USGS hydrologic unit runoff. We believe that this is the first time the USGS hydrologic unit runoff has been used in LSM evaluation and found to be very reasonable. (2) Groundwater evaluation indicated that Noah-MP captures the general spatial pattern of the climate conditions and captures the temporal patterns for wet regions. However, it fails in simulating the absolute values and the temporal variation in the water table for dry regions. (3) The addition of leaf dynamics to Noah-MP has improved its performance in ET simulation for natural land cover types. (4) One of the highlights of the study is that Noah-MP produces soil moisture values consistent with SCAN observations for the top two soil layers (0–10 cm and 10–40 cm), which indicates its great potential for use in studying land–atmosphere coupling. (5) The Noah-MP-simulated TWS anomaly agrees very well with GRACE observations, which may partly benefit from groundwater dynamics' inclusion in the model, considered the second largest component of the TWS anomaly for most of the MRB.

Chapter 3 used the NLDAS test bed to assess the hydrological performance of four LSMs: Noah, Noah-MP, CLM4, and VIC. All models were driven by the same NLDAS-2 atmospheric forcing and evaluated against the same observational data sets. Compared to Noah, the other three models show significant improvements in TWS and streamflow and moderate improvements in ET and soil moisture. Among these LSMs, Noah-MP shows the best performance in simulating soil moisture and is among the best in simulating TWS; CLM4 shows the best performance in simulating ET; and VIC shows the best performance in simulating streamflow.

118

Deficiencies are also found in these LSMs. The Noah-MP simulated ET grows too fast in the spring, which coincides with the fact that its modeled LAI peaks too soon. Although the CLM4 model produced the highest correlation coefficient for TWS anomaly for the entire CONUS, it produced either too high or too low of an amplitude of the annual TWS variation. In addition, CLM4 produced much weaker soil moisture variability than both SCAN observations and other LSMs. Finally, the VIC model consistently overestimated ET compared to observations from MODIS and FLUXNET.

In summary, this study gave an overview of the performances of four LSMs over CONUS. We demonstrated that, by providing reliable atmospheric forcing data, four in-house LSMs (with more to be added), most appropriate observational data, and necessary tools, the NLDAS test bed is a valid platform for evaluating land models on continental or large river basin scales in the U.S.

Chapter 4 described the coupling of N dynamics into Noah-MP by incorporating FUN's strength in plant N uptake and SWAT's strength in soil N cycling and agricultural management. The new model was then evaluated at KBS where provides good-quality, long-term observed N and ecological data. The model-simulated soil moisture is consistent with observation, which shows that Noah-MP provides a good water environment for the N cycling. The simulated concentrations of soil nitrate and N leaching from soil bottom also compare well with observations. Although the model does not most effectively simulate some peaks, especially regarding N leaching, the averages are very close to the observed values and the correlation coefficients are reasonable. Considering the wide spread of the range error bars defined by the measurements at the six replicates, the model performs with high skill in capturing the major N flux/state variables. The significant improvement of annual NPP simulation demonstrates that the N limitation effect on plant growth is well represented in the model.

Moreover, the addition of N dynamics in Noah-MP improves the modeling of the C and water cycles. Comparing to the default Noah-MP, NPP simulations are improved significantly in terms of consistent averages and much higher correlation coefficients with observation. The temporal pattern of simulated ET is also improved, featuring lower ET during spring and delayed peak. This enhancement is expected to facilitate the simultaneous predictions of weather and environment by using a fully coupled system.

5.2 CONTRIBUTIONS TO THE UNDERSTANDING OF THE EARTH SYSTEM

My studies contribute to the continuing improvement of our understanding and modeling of climate and environmental processes by identifying the strengths and limitations of the state-of-the-science LSMs. For example, Noah-MP has shown high skills in simulating the major hydrological variables. The inclusion of groundwater model and dynamic leaf model improves the model's performances and enables model's capability in representing these dynamics. However, they are still in their infancy and need further improvement. Overall, Noah-MP and other LSMs are doing reasonable jobs in representing the natural processes; however, human related processes are still poorly represented in LSMs, which should be the direction for LSM development.

In line with this, by integrating the N dynamics into Noah-MP, I have enabled the model's capability of representing agricultural managements such as N fertilization and tillage. This provides a useful tool for understanding the interactions of the N cycle with climate and carbon and water cycles. The evaluation at the KBS site shows that the coupled model is capable of representing these interactions. The accurate characterization of N limitation on plant growth significantly improves C simulation. The inclusion of N dynamics decreases NPP by 20.6%, LAI by 3.2%, GPP by 2.6%, and NEE (sink) by 2.2%, but increases transpiration by 15.0%. All these C and water related variables are delayed

for about half month, which is considered closer to reality. Because these phenomena are from a single site only, they may not apply to other locations. Nonetheless, by providing a powerful modeling tool, this is a big step towards the goal of better understanding the interactions of climate, hydrological cycle, and biogeochemical cycles.

My studies explain some phenomena about TWS. Based on hydrological modeling using Noah-MP and CLM4, the contributions to TWS anomaly from soil moisture, groundwater, and snow are equally important for CONUS, with each contributing approximately 1/3. However, this largely depends on location, especially for snow. If we look at the state of Florida only, the contribution from snow is negligible. However, it may not be true that less snow makes lower contribution to TWS anomaly. For example, if we look at the Southeastern region, snow contributes less than 5% to the TWS anomaly in terms of R^2; however, it contributes more than 40% to the TWS anomaly in terms of $RMSE$. This is because in this region, snowfall varies largely from year to year. If we analyze the contributions of TWS anomaly from the water flux terms (precipitation, ET, and runoff) perspective, ET is the dominant flux. However, these may be model specific and we need to confirm this by collecting observations for all of these storage terms simultaneously for the same region.

5.3 FUTURE WORK

To address the uncertainties and limitations of this dissertation, future work needs to be done. First of all, as recommended in Chapter 2, the dynamic leaf scheme in Noah-MP can capture the LAI growth in natural land but not in cropland. In Chapter 4, the new Noah-MP with N dynamics has difficulty with reproducing observed annual NPP for some of the simulation years because currently the model is not capable of simulating specific crops. Both issues require a crop model to be included in Noah-MP in the future to simulate

specific crops. Other agricultural management practices also need to be considered, such as tillage, harvest, and tile drainage.

River routing—horizontal water transport between grid cells—is considered in Chapter 3 but neglected in Chapter 2; this is an issue often neglected in coupled weather and climate models. This may not be a big issue in early weather and climate models at coarse resolution. With the increasing of model resolution (Wood et al., 2011), however, the effect of horizontal water transport becomes more pronounced on weather and climate predictions and thus the inclusion of this process becomes an urgent task.

In Chapter 4, since the N dynamic model is newly coupled into Noah-MP and only evaluated at one experimental site, we need to make more of an effort in model evaluation and improvement. The logical next step is to evaluate the model at additional sites and expand evaluation to regional and/or global scales. To make the model capable of doing water quality modeling, a riverine N transport scheme needs to be included. Moreover, soil organic matter dynamics affects the organic N storage and hence the decomposition and mineralization rates. Therefore, we need to improve the parameterization for soil organic matter dynamics. Finally, since phosphorus co-limits plant growth (Elser et al., 2007) and is also an important nutrient causing eutrophication (Carpenter, 2008; Schindler et al., 2008; Conley et al., 2009), Noah-MP needs to include a phosphorus dynamics model.

In summary, the studies presented here not only identified the strengths and limitations of Noah-MP and the other three LSMs but also significantly advanced the model development of Noah-MP by completing its plant and soil N dynamics. I hope these can benefit future LSM development. The ultimate goal of these studies is to facilitate the application of LSMs in climate change mitigation and environmental protection.

Appendix: Acronyms

BNF	Biological nitrogen fixation
C	Carbon
CLM	Community Land Model
CONUS	The continental United States
ET	Evapotranspiration
FACE	Free Air CO2 Enrichment
FUN	Fixation and Uptake of Nitrogen
GLASS	Global Land/Atmosphere System Study
GPP	Gross primary production
GRACE	Gravity Recovery and Climate Experiment
GVF	Green vegetation fraction
IDS	Interdisciplinary Research in Earth Science
KBS	Kellogg Biological Station
LAI	Leaf area index
LSM	Land Surface Model
LTER	Long-term Ecological Research
MODIS	Moderate-resolution Imaging Spectroradiometer
MRB	Mississippi River Basin
N	Nitrogen
NCEP	NOAA's National Centers for Environmental Prediction
NEE	Net ecosystem exchange
NLDAS	North American Land Data Assimilation System
NOAA	National Oceanic and Atmospheric Administration

Noah-MP	Noah Land Surface Model with MultiParameterization options
NPP	Net primary productivity
NSE	Nash–Sutcliffe efficiency
NWS	National Weather Service
PRISM	Parameter–elevation Regressions on Independent Slopes Model
RMSE	Root-mean-square error
SCAN	USDA Soil Climate Analysis Network
SIMGM	SIMple Groundwater Model
SMC	Soil moisture
SOM	Soil organic matters
STATSGO	State Soil Geographic Database
SWAT	Soil and Water Assessment Tool
SWE	Snow water equivalent
TWS	Terrestrial water storage
UNH-GRDC	The University of New Hampshire/Global Runoff Data Centre
USDA	United States Department of Agriculture
USGS	United States Geological Survey
USLE	Universal soil loss equation
VIC	Variable Infiltration Capacity
WRF	Weather Research and Forecasting
WTD	Water table depth

References

Abramowitz, G., 2012: Towards a public, standardized, diagnostic benchmarking system for land surface models. *Geosci. Model Dev.*, **5**(3): 819-827, doi:10.5194/gmd-5-819-2012.

Alexeev, V.A., D.J. Nicolsky, V.E. Romanovsky and D.M. Lawrence, 2007: An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost. *Geophys. Res. Lett.*, **34**(9), doi:10.1029/2007GL029536.

Algoazany, A.S., P.K. Kalita, G.F. Czapar and J.K. Mitchell, 2007: Phosphorus transport through subsurface drainage and surface runoff from a flat watershed in east central Illinois, USA. *J. Environ. Qual.*, **36**(3): 681-693, doi:10.2134/Jeq2006.0161.

Bachelet, D., J.M. Lenihan, C. Daly, R.P. Neilson, D.S. Ojima and W.J. Parton, 2001: MC1: A dynamic vegetation model for estimating the distribution of vegetation and associated ecosystem fluxes of carbon, nutrients, and water: Technical documentation version 1.0. *General Technical Report PNW-GTR-508*, U.S. Dept. of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA, 95 pp.

Baldocchi, D., E. Falge, L.H. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R. Evans 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. Amer. Meteor. Soc.*, **82**(11): 2415-2434, doi:10.1175/1520-0477(2001)082<2415:Fantts>2.3.Co;2.

Barlage, M., M. Tewari, F. Chen, G. Miguez-Macho, Z.-L. Yang and G.-Y. Niu, 2015: The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP. *Clim. Change*, **129**(3-4): 485-498, doi:10.1007/s10584-014-1308-8.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle Jr, A.S. Donigian Jr and R.C. Johanson, 1997: Hydrological Simulation Program–FORTRAN, User's manual for release 11. *EPA/600/R-97/080*, U.S. Environmental Protection Agency, Athens, GA, 755 pp.

Blyth, E., J. Gash, A. Lloyd, M. Pryor, G.P. Weedon and J. Shuttleworth, 2010: Evaluating the JULES land surface model energy fluxes using FLUXNET data. *J. Hydrometeorol.*, **11**(2): 509-519, doi:10.1175/2009jhm1183.1.

Boesch, D.F., W.R. Boynton, L.B. Crowder, R.J. Diaz, R.W. Howarth, L.D. Mee, S.W. Nixon, N.N. Rabalais, R. Rosenberg, J.G. Sanders et al., 2009: Nutrient enrichment drives Gulf of Mexico hypoxia. *Eos Trans. AGU*, **90**(14): 117, doi:10.1029/2009eo140001.

Bonan, G.B., 1991: Atmosphere-biosphere exchange of carbon dioxide in boreal forests. *J. Geophys. Res.*, **96**(D4): 7301-7312, doi:10.1029/90jd02713.

Bonan, G.B., 1995: Land-atmosphere CO2 exchange simulated by a land surface process model coupled to an atmospheric general circulation model. *J. Geophys. Res.*, **100**(D2): 2817-2831.

Bonan, G.B., 2008: *Ecological Climatology: Concepts and Applications*. 2nd ed. Cambridge University Press, Cambridge, UK, 550 pp.

Bonan, G.B., P.J. Lawrence, K.W. Oleson, S. Levis, M. Jung, M. Reichstein, D.M. Lawrence and S.C. Swenson, 2011: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *J. Geophys. Res.*, **116**, doi:10.1029/2010jg001593.

Bonan, G.B. and S. Levis, 2010: Quantifying carbon-nitrogen feedbacks in the Community Land Model (CLM4). *Geophys. Res. Lett.*, **37**, doi:10.1029/2010gl042430.

Boyer, E.W., R.W. Howarth, J.N. Galloway, F.J. Dentener, P.A. Green and C.J. Vorosmarty, 2006: Riverine nitrogen export from the continents to the coasts. *Global Biogeochem. Cy.*, **20**(1), doi:10.1029/2005gb002537.

Brakebill, J.W., D.M. Wolock and S.E. Terziotti, 2011: Digital hydrologic networks supporting applications related to spatially referenced regression modeling. *J. Am. Water Resour. Assoc.*, **47**(5): 916-932, doi:10.1111/j.1752-1688.2011.00578.x.

Brandt, M. and H. Ejhed, 2002: Transport-Retention-Kallfordelning, Belastning pa havet. *Report 5247*, Swedish Environmental Protection Agency, Lindblom & Co., Stockholm, Sweden.

Brzostek, E.R., J.B. Fisher and R.P. Phillips, 2014: Modeling the carbon cost of plant nitrogen acquisition: Mycorrhizal trade-offs and multipath resistance uptake improve predictions of retranslocation. *J. Geophys. Res.*, **119**(8): 1684-1697, doi:10.1002/2014jg002660.

Burnash, R.J.C., R.L. Ferral and R.A. McGuire, 1973: A generalized streamflow simulation system: Conceptual modeling for digital computers, Joint Federal-State River Forecast Center, Sacramento, CA, USA, 204 pp.

Cai, X., Z.-L. Yang, C.H. David, G.-Y. Niu and M. Rodell, 2014a: Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin. *J. Geophys. Res.*, **119**(1): 23-38, doi:10.1002/2013jd020792.

Cai, X., Z.-L. Yang, Y. Xia, M. Huang, H. Wei, L.R. Leung and M.B. Ek, 2014b: Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over

CONUS using the NLDAS test bed. *J. Geophys. Res.*, **119**(24): 13751-13770, doi:10.1002/2014jd022113.

Carpenter, S.R., 2008: Phosphorus control is critical to mitigating eutrophication. *P. Natl. Acad. Sci. USA*, **105**(32): 11039-11040, doi:10.1073/pnas.0806112105.

Chen, F., Z. Janjic and K. Mitchell, 1997: Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Bound.-Layer Meteorol.*, **85**(3): 391-421, doi:10.1023/A:1000531001463.

Chen, J.L., C.R. Wilson and K.W. Seo, 2006: Optimized smoothing of gravity recovery and climate experiment (GRACE) time-variable gravity observations. *J. Geophys. Res.*, **111**(B6), doi:10.1029/2005jb004064.

Chen, Y., I. Velicogna, J.S. Famiglietti and J.T. Randerson, 2013: Satellite observations of terrestrial water storage provide early warning information about drought and fire season severity in the Amazon. *J. Geophys. Res.*, **118**(2): 495-504, doi:10.1002/Jgrg.20046.

Cherkauer, K.A., L.C. Bowling and D.P. Lettenmaier, 2003: Variable infiltration capacity cold land process model updates. *Global Planet. Change*, **38**(1-2): 151-159, doi:10.1016/S0921-8181(03)00025-0.

Christensen, N.S., A.W. Wood, N. Voisin, D.P. Lettenmaier and R.N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Clim. Change*, **62**(1-3): 337-363, doi:10.1023/B:Clim.0000013684.13621.1f.

Clark, D.B., L.M. Mercado, S. Sitch, C.D. Jones, N. Gedney, M.J. Best, M. Pryor, G.G. Rooney, R.L.H. Essery, E. Blyth et al., 2011a: The Joint UK Land Environment Simulator (JULES), model description - Part 2: Carbon fluxes and vegetation dynamics. *Geosci. Model Dev.*, **4**(3): 701-722, doi:10.5194/gmd-4-701-2011.

Clark, M.P., D. Kavetski and F. Fenicia, 2011b: Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resour. Res.*, **47**, doi:10.1029/2010wr009827.

Collins, W.D., C.M. Bitz, M.L. Blackmon, G.B. Bonan, C.S. Bretherton, J.A. Carton, P. Chang, S.C. Doney, J.J. Hack, T.B. Henderson et al., 2006: The Community Climate System Model version 3 (CCSM3). *J. Climate*, **19**(11): 2122-2143, doi:10.1175/Jcli3761.1.

127

Conley, D.J., H.W. Paerl, R.W. Howarth, D.F. Boesch, S.P. Seitzinger, K.E. Havens, C. Lancelot and G.E. Likens, 2009: Controlling eutrophication: Nitrogen and phosphorus. *Science*, **323**(5917): 1014-1015, doi:10.1126/science.1167755.

Cosby, B.J., G.M. Hornberger, R.B. Clapp and T.R. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.*, **20**(6): 682-690, doi:10.1029/Wr020i006p00682.

Cosgrove, B.A., D. Lohmann, K.E. Mitchell, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, C. Marshall, J. Sheffield, Q.Y. Duan et al., 2003a: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res.*, **108**(D22), doi:10.1029/2002jd003118.

Cosgrove, B.A., D. Lohmann, K.E. Mitchell, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, J. Sheffield, Q.Y. Duan, L.F. Luo et al., 2003b: Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS). *J. Geophys. Res.*, **108**(D22), doi:10.1029/2002jd003316.

Cox, P.M., R.A. Betts, C.D. Jones, S.A. Spall and I.J. Totterdell, 2000: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**(6809): 184-187, doi:10.1038/35041539.

Cramer, W., A. Bondeau, F.I. Woodward, I.C. Prentice, R.A. Betts, V. Brovkin, P.M. Cox, V. Fisher, J.A. Foley, A.D. Friend et al., 2001: Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models. *Global Change Biol.*, **7**(4): 357-373, doi:10.1046/j.1365-2486.2001.00383.x.

Daly, C., R.P. Neilson and D.L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**(2): 140-158, doi:10.1175/1520-0450(1994)033<0140:Astmfm>2.0.Co;2.

Diaz, R.J. and R. Rosenberg, 2008: Spreading dead zones and consequences for marine ecosystems. *Science*, **321**(5891): 926-929, doi:10.1126/science.1156401.

Dickinson, R., A. Henderson-Sellers, P. Kennedy and M. Wilson, 1986: Biosphere-atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model. *NCAR/TN-275+STR*, University Corporation for Atmospheric Research, Boulder, CO, 69 pp.

Dickinson, R.E., J.A. Berry, G.B. Bonan, G.J. Collatz, C.B. Field, I.Y. Fung, M. Goulden, W.A. Hoffmann, R.B. Jackson, R. Myneni et al., 2002: Nitrogen controls on climate model evapotranspiration. *J. Climate*, **15**(3): 278-295, doi:10.1175/1520-0442(2002)015<0278:Ncocme>2.0.Co;2.

Dickinson, R.E., K.W. Oleson, G. Bonan, F. Hoffman, P. Thornton, M. Vertenstein, Z.-L. Yang and X.B. Zeng, 2006: The Community Land Model and its climate statistics as a component of the Community Climate System Model. *J. Climate*, **19**(11): 2302-2324, doi:10.1175/Jcli3742.1.

Dickinson, R.E., M. Shaikh, R. Bryant and L. Graumlich, 1998: Interactive canopies for a climate model. *J. Climate*, **11**(11): 2823-2836, doi:10.1175/1520-0442(1998)011<2823:Icfacm>2.0.Co;2.

Donner, S.D. and D. Scavia, 2007: How climate controls the flux of nitrogen by the Mississippi River and the development of hypoxia in the Gulf of Mexico. *Limnol. Oceanogr.*, **52**(2): 856-861.

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), doi:10.1029/2002jd003296.

Ek, M.B., Y. Xia, E. Wood, J. Sheffield, L. Luo, D. Lettenmaier, B. Livneh, D. Mocko, B. Cosgrove, J. Meng et al., 2011: North American Land Data Assimilation System Phase 2 (NLDAS-2): Development and applications. *GEWEX News*, **21**(2): 6-7.

Elser, J.J., M.E.S. Bracken, E.E. Cleland, D.S. Gruner, W.S. Harpole, H. Hillebrand, J.T. Ngai, E.W. Seabloom, J.B. Shurin and J.E. Smith, 2007: Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.*, **10**(12): 1135-1142, doi:10.1111/j.1461-0248.2007.01113.x.

Entin, J.K., A. Robock, K.Y. Vinnikov, S.E. Hollinger, S.X. Liu and A. Namkhai, 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res.*, **105**(D9): 11865-11877, doi:10.1029/2000jd900051.

Falloon, P., R. Betts, A. Wiltshire, R. Dankers, C. Mathison, D. McNeall, P. Bates and M. Trigg, 2011: Validation of river flows in HadGEM1 and HadCM3 with the TRIP river flow model. *J. Hydrometeorol.*, **12**(6): 1157-1180, doi:10.1175/2011jhm1388.1.

Famiglietti, J.S., M. Lo, S.L. Ho, J. Bethune, K.J. Anderson, T.H. Syed, S.C. Swenson, C.R. de Linage and M. Rodell, 2011: Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophys. Res. Lett.*, **38**, doi:10.1029/2010gl046442.

Fan, Y. and G. Miguez-Macho, 2011: A simple hydrologic framework for simulating wetlands in climate and earth system models. *Clim. Dyn.*, **37**(1-2): 253-278, doi:10.1007/s00382-010-0829-8.

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**, doi:10.1029/2006jd008111.

Fekete, B.M., C.J. Vorosmarty and W. Grabs, 2002: High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biogeochem. Cy.*, **16**(3), doi:10.1029/1999gb001254.

Fisher, J.B., G. Badgley and E. Blyth, 2012: Global nutrient limitation in terrestrial vegetation. *Global Biogeochem. Cy.*, **26**, doi:10.1029/2011gb004252.

Fisher, J.B., S. Sitch, Y. Malhi, R.A. Fisher, C. Huntingford and S.Y. Tan, 2010: Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation. *Global Biogeochem. Cy.*, **24**, doi:10.1029/2009gb003621.

Flanner, M.G., C.S. Zender, J.T. Randerson and P.J. Rasch, 2007: Present-day climate forcing and response from black carbon in snow. *J. Geophys. Res.*, **112**(D11): D11202, doi:10.1029/2006jd008003.

Foley, J.A., I.C. Prentice, N. Ramankutty, S. Levis, D. Pollard, S. Sitch and A. Haxeltine, 1996: An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochem. Cy.*, **10**(4): 603-628.

Gao, H.L., Q.H. Tang, C.R. Ferguson, E.F. Wood and D.P. Lettenmaier, 2010: Estimating the water budget of major US river basins via remote sensing. *Int. J. Remote Sens.*, **31**(14): 3955-3978, doi:10.1080/01431161.2010.483488.

Gent, P.R., S.G. Yeager, R.B. Neale, S. Levis and D.A. Bailey, 2010: Improvements in a half degree atmosphere/land version of the CCSM. *Clim. Dyn.*, **34**(6): 819-833, doi:10.1007/s00382-009-0614-8.

Gentry, L.E., M.B. David, F.E. Below, T.V. Royer and G.F. Mclsaac, 2009: Nitrogen mass balance of a tile-drained agricultural watershed in east-central Illinois. *J. Environ. Qual.*, **38**(5): 1841-1847, doi:10.2134/Jeq2008.0406.

Goswami, D., 2006: Analysis of flow and water quality components in drainage channels in tile drained watersheds. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Illinois, USA, 154 pp.

Gruber, N. and J.N. Galloway, 2008: An Earth-system perspective of the global nitrogen cycle. *Nature*, **451**(7176): 293-296, doi:10.1038/Nature06592.

Güntner, A., 2008: Improvement of global hydrological models using GRACE data. *Surv. Geophys.*, **29**(4-5): 375-397, doi:10.1007/s10712-008-9038-y.

Guo, Z.C. and P.A. Dirmeyer, 2006: Evaluation of the Second Global Soil Wetness Project soil moisture simulations: 1. Intermodel comparison. *J. Geophys. Res.*, **111**(D22), doi:10.1029/2006jd007233.

Gutman, G. and A. Ignatov, 1998: The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.*, **19**(8): 1533-1543, doi:10.1080/014311698215333.

Gutowski, W.J., C.J. Vorosmarty, M. Person, Z. Otles, B. Fekete and J. York, 2002: A coupled land-atmosphere simulation program (CLASP): Calibration and validation. *J. Geophys. Res.*, **107**(D16), doi:10.1029/2001jd000392.

Haith, D.A. and L.L. Shoemaker, 1987: Generalized watershed loading functions for stream flow nutrients. *Water Resour. Bull.*, **23**(3): 471-478.

Hogue, T.S., L.A. Bastidas, H.V. Gupta and S. Sorooshian, 2006: Evaluating model performance and parameter behavior for varying levels of land surface model complexity. *Water Resour. Res.*, **42**(8), doi:10.1029/2005wr004440.

Hou, Z., M. Huang, L.R. Leung, G. Lin and D.M. Ricciuto, 2012: Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model. *J. Geophys. Res.*, **117**(D15): D15108, doi:10.1029/2012jd017521.

Houborg, R., M. Rodell, B.L. Li, R. Reichle and B.F. Zaitchik, 2012: Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resour. Res.*, **48**, doi:10.1029/2011wr011291.

Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der LINDEN, X. Dai and K. Maskell, 2001: *Climate change 2001: The scientific basis: Contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge; New York, x, 881 pp.

Houlton, B.Z., Y.P. Wang, P.M. Vitousek and C.B. Field, 2008: A unifying framework for dinitrogen fixation in the terrestrial biosphere. *Nature*, **454**(7202): 327-U34, doi:10.1038/Nature07028.

Howarth, R.W., D.P. Swaney, E.W. Boyer, R. Marino, N. Jaworski and C. Goodale, 2006: The influence of climate on average nitrogen export from large watersheds in the Northeastern United States. *Biogeochemistry*, **79**(1-2): 163-186, doi:10.1007/s10533-006-9010-1.

Huang, M., Z. Hou, L.R. Leung, Y. Ke, Y. Liu, Z. Fang and Y. Sun, 2013: Uncertainty analysis of runoff simulations and parameter identifiability in the Community Land Model: Evidence from MOPEX basins. *J. Hydrometeorol.*, **14**(6): 1754-1772, doi:10.1175/JHM-D-12-0138.1.

Hungate, B.A., J.S. Dukes, M.R. Shaw, Y.Q. Luo and C.B. Field, 2003: Nitrogen and climate change. *Science*, **302**(5650): 1512-1513, doi:10.1126/science.1091390.

Jiang, X.Y., G.Y. Niu and Z.L. Yang, 2009: Impacts of vegetation and groundwater dynamics on warm season precipitation over the Central United States. *J. Geophys. Res.*, **114**, doi:10.1029/2008jd010756.

Jordan, R., 1991: A One-dimensional Temperature Model for a Snow Cover: Technical Documentation for SNTHERM.89. *Special Report SR+91-16*, Cold Regions Research and Engineering Laboratory, U.S. Army Corps of Engineers, Hanover, New Hampshire, 49 pp.

Jung, M., M. Reichstein and A. Bondeau, 2009: Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, **6**(10): 2001-2013.

Jung, M., M. Reichstein, P. Ciais, S.I. Seneviratne, J. Sheffield, M.L. Goulden, G. Bonan, A. Cescatti, J.Q. Chen, R. de Jeu et al., 2010: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, **467**(7318): 951-954, doi:10.1038/Nature09396.

Ke, Y., L.R. Leung, M. Huang, A.M. Coleman, H. Li and M.S. Wigmosta, 2012: Development of high resolution land surface parameters for the Community Land Model. *Geosci. Model Dev.*, **5**(6): 1341-1362, doi:10.5194/gmd-5-1341-2012.

Kicklighter, D.W., M. Bruno, S. Donges, G. Esser, M. Heimann, J. Helfrich, F. Ift, F. Joos, J. Kaduk, G.H. Kohlmaier et al., 1999: A first-order analysis of the potential role of CO_2 fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models. *Tellus B*, **51**(2): 343-366, doi:10.1034/j.1600-0889.1999.00017.x.

Knops, J.M.H. and D. Tilman, 2000: Dynamics of soil nitrogen and carbon accumulation for 61 years after agricultural abandonment. *Ecology*, **81**(1): 88-98, doi:10.1890/0012-9658(2000)081[0088:Dosnac]2.0.Co;2.

Kollet, S.J. and R.M. Maxwell, 2008: Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resour. Res.*, **44**(2), doi:10.1029/2007wr006004.

Komatsu, E., T. Fukushima and H. Harasawa, 2007: A modeling approach to forecast the effect of long-term climate change on lake water quality. *Ecol. Model.*, **209**(2-4): 351-366, doi:10.1016/j.ecolmodel.2007.07.021.

Koster, R.D. and M.J. Suarez, 1996: Energy and water balance calculations in the Mosaic LSM. *NASA TM-104606*, Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, MD, USA, 60 pp.

Kronvang, B., H. Behrendt, H.E. Andersen, B. Arheimer, A. Barr, S.A. Borgvang, F. Bouraoui, K. Granlund, B. Grizzetti, P. Groenendijk et al., 2009: Ensemble modelling of nutrient loads and nutrient load partitioning in 17 European catchments. *J. Environ. Monitor.*, **11**(3): 572-583, doi:10.1039/B900101h.

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**, doi:10.1029/2010jd014909.

Kumar, S.V., C.D. Peters-Lidard, J. Santanello, K. Harrison, Y. Liu and M. Shaw, 2012: Land surface Verification Toolkit (LVT) - A generalized framework for land surface model evaluation. *Geosci. Model Dev.*, **5**(3): 869-886, doi:10.5194/gmd-5-869-2012.

Landerer, F.W. and S.C. Swenson, 2012: Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour. Res.*, **48**, doi:10.1029/2011wr011453.

Lawrence, D.M., K.W. Oleson, M.G. Flanner, C.G. Fletcher, P.J. Lawrence, S. Levis, S.C. Swenson and G.B. Bonan, 2012: The CCSM4 land simulations, 1850–2005: Assessment of surface climate and new capabilities. *J. Climate*, **25**(7): 2240-2260, doi:10.1175/jcli-d-11-00103.1.

Lawrence, D.M., K.W. Oleson, M.G. Flanner, P.E. Thornton, S.C. Swenson, P.J. Lawrence, X.B. Zeng, Z.-L. Yang, S. Levis, K. Sakaguchi et al., 2011: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *J. Adv. Model. Earth Sy.*, **3**, doi:10.1029/2011ms000045.

Lawrence, D.M. and A.G. Slater, 2008: Incorporating organic soil into a global climate model. *Clim. Dyn.*, **30**(2-3): 145-160, doi:10.1007/s00382-007-0278-1.

Lei, H., M. Huang, L.R. Leung, D. Yang, X. Shi, J. Mao, D.J. Hayes, C.R. Schwalm, Y. Wei and S. Liu, 2014: Sensitivity of global terrestrial gross primary production to hydrologic states simulated by the Community Land Model using two runoff parameterizations. *J. Adv. Model. Earth Sy.*, doi:10.1002/2013MS000252.

Leng, G., M. Huang, Q. Tang, H. Gao and L.R. Leung, 2013: Modeling the effects of groundwater-fed irrigation on terrestrial hydrology over the conterminous United States. *J. Hydrometeorol.*, doi:10.1175/JHM-D-13-049.1.

Leung, L.R., M.Y. Huang, Y. Qian and X. Liang, 2011: Climate-soil-vegetation control on groundwater table dynamics and its feedbacks in a climate model. *Clim. Dyn.*, **36**(1-2): 57-81, doi:10.1007/s00382-010-0746-x.

Li, H., M. Sivapalan, F. Tian and D. Liu, 2010: Water and nutrient balances in a large tile-drained agricultural catchment: a distributed modeling study. *Hydrol. Earth Syst. Sci.*, **14**(11): 2259-2275, doi:10.5194/hess-14-2259-2010.

Li, H.Y., M.Y. Huang, M.S. Wigmosta, Y.H. Ke, A.M. Coleman, L.R. Leung, A.H. Wang and D.M. Ricciuto, 2011: Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed. *J. Geophys. Res.*, **116**, doi:10.1029/2011jd016276.

Liang, X., D.P. Lettenmaier, E.F. Wood and S.J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.*, **99**(D7): 14415-14428.

Liu, Q., R.H. Reichle, R. Bindlish, M.H. Cosh, W.T. Crow, R. de Jeu, G.J.M. De Lannoy, G.J. Huffman and T.J. Jackson, 2011: The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *J. Hydrometeorol.*, **12**(5): 750-765, doi:10.1175/Jhm-D-10-05000.1.

Livneh, B., E.A. Rosenberg, C.Y. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer and D.P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate*, **26**(23): 9384-9392, doi:10.1175/Jcli-D-12-00508.1.

Livneh, B., Y.L. Xia, K.E. Mitchell, M.B. Ek and D.P. Lettenmaier, 2010: Noah LSM snow model diagnostics and enhancements. *J. Hydrometeorol.*, **11**(3): 721-738, doi:10.1175/2009jhm1174.1.

Lo, M.H., J.S. Famiglietti, P.J.F. Yeh and T.H. Syed, 2010: Improving parameter estimation and water table depth simulation in a land surface model using GRACE

water storage and estimated base flow data. *Water Resour. Res.*, **46**, doi:10.1029/2009wr007855.

Lohmann, D., K.E. Mitchell, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, B.A. Cosgrove, J. Sheffield, Q.Y. Duan, L.F. Luo et al., 2004: Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project. *J. Geophys. Res.*, **109**(D7), doi:10.1029/2003jd003517.

Lohmann, D., E. Raschke, B. Nijssen and D.P. Lettenmaier, 1998: Regional scale hydrology: II. Application of the VIC-2L model to the Weser River, Germany. *Hydrolog. Sci. J.*, **43**(1): 143-158, doi:10.1080/02626669809492108.

Long, D., B.R. Scanlon, L. Longuevergne, A.Y. Sun, D.N. Fernando and H. Save, 2013: GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. *Geophys. Res. Lett.*, **40**(13): 3395-3401, doi:10.1002/Grl.50655.

Luo, Y., B. Su, W.S. Currie, J.S. Dukes, A.C. Finzi, U. Hartwig, B. Hungate, R.E. McMurtrie, R. Oren, W.J. Parton et al., 2004: Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. *Bioscience*, **54**(8): 731-739, doi:10.1641/0006-3568(2004)054[0731:Pnloer]2.0.Co;2.

Luo, Y.Q., J.T. Randerson, G. Abramowitz, C. Bacour, E. Blyth, N. Carvalhais, P. Ciais, D. Dalmonech, J.B. Fisher, R. Fisher et al., 2012: A framework for benchmarking land models. *Biogeosciences*, **9**(10): 3857-3874, doi:10.5194/bg-9-3857-2012.

Manabe, S., 1969: Climate and ocean circulation .I. Atmospheric circulation and hydrology of earths surface. *Mon. Wea. Rev.*, **97**(11): 739-774.

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. Climate*, **15**(22): 3237-3251, doi:10.1175/1520-0442(2002)015<3237:Althbd>2.0.Co;2.

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**(D13), doi:10.1029/2006jd008112.

Miller, D.A. and R.A. White, 1998: A Conterminous United States Multilayer Soil Characteristics Dataset for Regional Climate and Hydrology Modeling. *Earth Interact.*, **2**(2): 1-26, doi:10.1175/1087-3562(1998)002<0001:acusms>2.3.co;2.

Mitchell, K.E., D. Lohmann, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, B.A. Cosgrove, J. Sheffield, Q.Y. Duan, L.F. Luo et al., 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, **109**(D7), doi:10.1029/2003jd003823.

Monastersky, R., 2015: Anthropocene: The human age. *Nature*, **519**(7542): 144-147, doi:10.1038/519144a.

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel and T.L. Veith, 2007: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, **50**(3): 885-900.

Mu, Q., F.A. Heinsch, M. Zhao and S.W. Running, 2007: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens. Environ.*, **111**(4): 519-536, doi:10.1016/j.rse.2007.04.015.

Mu, Q.Z., M.S. Zhao and S.W. Running, 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.*, **115**(8): 1781-1800, doi:10.1016/j.rse.2011.02.019.

Murray, S.J., P.N. Foster and I.C. Prentice, 2011: Evaluation of global continental hydrology as simulated by the Land-surface Processes and eXchanges Dynamic Global Vegetation Model. *Hydrol. Earth Syst. Sci.*, **15**(1): 91-105, doi:10.5194/hess-15-91-2011.

Murray, S.J., I.M. Watson and I.C. Prentice, 2013: The use of dynamic global vegetation models for simulating hydrology and the potential integration of satellite observations. *Progress in Physical Geography*, **37**(1): 63-97, doi:10.1177/0309133312460072.

Nash, J.E. and J.V. Sutcliffe, 1970: River flow forecasting through conceptual models. Part 1. A discussion of principles. *J. Hydrol.*, **10**(3): 282-290, doi:10.1016/0022-1694(70)90255-6.

National Research Council, 2000: Clean coastal waters: understanding and reducing the effects of nutrient pollution, National Academy of Sciences, Washington, DC.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry and J.R. Williams, 2011: Soil and Water Assessment Tool theoretical documentation version 2009. *Technical Report No. 406*, Texas Water Resources Institute, Texas A&M University, College Station, TX.

Nicolsky, D.J., V.E. Romanovsky, V.A. Alexeev and D.M. Lawrence, 2007: Improved modeling of permafrost dynamics in a GCM land-surface scheme. *Geophys. Res. Lett.*, **34**(8): L08501, doi:10.1029/2007GL029525.

Nijssen, B., D.P. Lettenmaier, X. Liang, S.W. Wetzel and E.F. Wood, 1997: Streamflow simulation for continental-scale river basins. *Water Resour. Res.*, **33**(4): 711-724, doi:10.1029/96wr03517.

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**(5): 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**(D21), 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**(D7), doi:10.1029/2006jd007522.

Niu, G.-Y., Z.-L. Yang, K.E. Mitchell, F. Chen, M.B. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E. Rosero et al., 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, doi:10.1029/2010jd015139.

Oleson, K., D. Lawrence, G. Bonan, M. Flanner, E. Kluzek, P. Lawrence, S. Levis, S. Swenson, P. Thornton, A. Dai et al., 2010: Technical description of version 4.0 of the Community Land Model (CLM). *NCAR Technical Note NCAR/TN-478+STR*, University Corporation for Atmospheric Research, Boulder, CO, USA, 266 pp.

Oleson, K.W., D.M. Lawrence, G.B. Bonan, B. Drewniak, M. Huang, C.D. Koven, S. Levis, F. Li, W.J. Riley, Z.M. Subin et al., 2013: Technical description of version 4.5 of the Community Land Model (CLM). *NCAR Technical Note NCAR/TN-503+STR*, National Center for Atmospheric Research, Boulder, Colorado, 434 pp.

Oleson, K.W., G.-Y. Niu, Z.-L. Yang, D.M. Lawrence, P.E. Thornton, P.J. Lawrence, R. Stockli, R.E. Dickinson, G.B. Bonan, S. Levis et al., 2008: Improvements to the Community Land Model and their impact on the hydrological cycle. *J. Geophys. Res.*, **113**(G1), doi:10.1029/2007jg000563.

Ostle, N.J., P. Smith, R. Fisher, F.I. Woodward, J.B. Fisher, J.U. Smith, D. Galbraith, P. Levy, P. Meir, N.P. McNamara et al., 2009: Integrating plant-soil interactions into

global carbon cycle models. *J. Ecol.*, **97**(5): 851-863, doi:10.1111/j.1365-2745.2009.01547.x.

Ozdogan, M., M. Rodell, H.K. Beaudoing and D.L. Toll, 2010: Simulating the effects of irrigation over the United States in a land surface model based on satellite-derived agricultural data. *J. Hydrometeorol.*, **11**(1): 171-184, doi:10.1175/2009jhm1116.1.

Pitman, A.J., 2003: The evolution of, and revolution in, land surface schemes designed for climate models. *Int. J. Climatol.*, **23**(5): 479-510, doi:10.1002/Joc.893.

Pokhrel, Y., N. Hanasaki, S. Koirala, J. Cho, P.J.F. Yeh, H. Kim, S. Kanae and T. Oki, 2012: Incorporating anthropogenic water regulation modules into a land surface model. *J. Hydrometeorol.*, **13**(1): 255-269, doi:10.1175/Jhm-D-11-013.1.

Prince, S.D., J. Haskett, M. Steininger, H. Strand and R. Wright, 2001: Net primary production of US Midwest croplands from agricultural harvest yield data. *Ecol. Appl.*, **11**(4): 1194-1205, doi:10.2307/3061021.

Rasmussen, R., K. Ikeda, C.H. Liu, D. Gochis, M. Clark, A.G. Dai, E. Gutmann, J. Dudhia, F. Chen, M. Barlage et al., 2014: Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations. *J. Hydrometeorol.*, **15**(3): 1091-1116, doi:10.1175/Jhm-D-13-0118.1.

Roads, J., R. Lawford, E. Bainto, E. Berbery, S. Chen, B. Fekete, K. Gallo, A. Grundstein, W. Higgins, M. Kanamitsu et al., 2003: GCIP water and energy budget synthesis (WEBS). *J. Geophys. Res.*, **108**(D16), doi:10.1029/2002jd002583.

Rodell, M., J.L. Chen, H. Kato, J.S. Famiglietti, J. Nigro and C.R. Wilson, 2007: Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeol. J.*, **15**(1): 159-166, doi:10.1007/s10040-006-0103-7.

Rosero, E., Z.-L. Yang, L.E. Gulden, G.-Y. Niu and D.J. Gochis, 2009: Evaluating enhanced hydrological representations in Noah LSM over transition zones: Implications for model development. *J. Hydrometeorol.*, **10**(3): 600-622, doi:10.1175/2009jhm1029.1.

Rosero, E., Z.-L. Yang, T. Wagener, L.E. Gulden, S. Yatheendradas and G.-Y. Niu, 2010: Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season. *J. Geophys. Res.*, **115**, doi:10.1029/2009jd012035.

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**(1): 108-127, doi:10.1016/S0034-4257(99)00061-9.

Running, S.W. and J.C. Coughlan, 1988: A general model of forest ecosystem processes for regional applications I. Hydrologic balance, canopy. *Ecol. Model.*, **42**(2): 125-154, doi:10.1016/0304-3800(88)90112-3.

Scanlon, B.R., R.C. Reedy and K.F. Bronson, 2008: Impacts of land use change on nitrogen cycling archived in semiarid unsaturated zone nitrate profiles, southern High Plains, Texas. *Environ. Sci. Technol.*, **42**(20): 7566-7572, doi:10.1021/Es800792w.

Schaake, J.C., V.I. Koren, Q.Y. Duan, K. Mitchell and F. Chen, 1996: Simple water balance model for estimating runoff at different spatial and temporal scales. *J. Geophys. Res.*, **101**(D3): 7461-7475, doi:10.1029/95jd02892.

Schaefer, G.L., M.H. Cosh and T.J. Jackson, 2007: The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *J. Atmos. Oceanic Technol.*, **24**(12): 2073-2077, doi:10.1175/2007jtecha930.1.

Schindler, D.W., R.E. Hecky, D.L. Findlay, M.P. Stainton, B.R. Parker, M.J. Paterson, K.G. Beaty, M. Lyng and S.E.M. Kasian, 2008: Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. *P. Natl. Acad. Sci. USA*, **105**(32): 11254-11258, doi:10.1073/pnas.0805108105.

Schoumans, O., C. Siderius and P. Groenendijk, 2009a: NL-CAT application to six European catchments. *Alterra report 1205*, Alterra, Wageningen, the Netherlands, 276 pp.

Schoumans, O.F., M. Silgram, P. Groenendijk, F. Bouraoui, H.E. Andersen, B. Kronvang, H. Behrendt, B. Arheimer, H. Johnsson, Y. Panagopoulos et al., 2009b: Description of nine nutrient loss models: capabilities and suitability based on their characteristics. *J. Environ. Monitor.*, **11**(3): 506-514, doi:10.1039/B823239c.

Schwarz, G.E., A. Hoos, R. Alexander and R. Smith, 2006: The SPARROW surface water-quality model: Theory, application, and user documentation. *Techniques and Methods 6-B3*, U.S. Geological Survey, Reston, VA, 248 pp.

Sellers, P.J., M.D. Heiser and F.G. Hall, 1992: Relations between surface conductance and spectral vegetation indices at intermediate (100 m^2 to 15 km^2) length scales. *J. Geophys. Res.*, **97**(D17): 19033-19059, doi:10.1029/92JD01096.

Sellers, P.J., Y. Mintz, Y.C. Sud and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.*, **43**(6): 505-531.

Sellers, P.J., D.A. Randall, G.J. Collatz, J.A. Berry, C.B. Field, D.A. Dazlich, C. Zhang, G.D. Collelo and L. Bounoua, 1996: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *J. Climate*, **9**(4): 676-705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2.

Sheffield, J., G. Goteti and E.F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate*, **19**(13): 3088-3111, doi:10.1175/Jcli3790.1.

Sitch, S., B. Smith, I.C. Prentice, A. Arneth, A. Bondeau, W. Cramer, J.O. Kaplan, S. Levis, W. Lucht, M.T. Sykes et al., 2003: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biol.*, **9**(2): 161-185.

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, M. Barker, K.G. Duda, Y. Huang, W. Wang and J.G. Powers, 2008: A description of the Advanced Research WRF version 3. *Technical Note: NCAR/TN-475+STR*, National Center for Atmospheric Research, Boulder, CO, USA, 113 pp.

Slater, A.G., T.J. Bohn, J.L. McCreight, M.C. Serreze and D.P. Lettenmaier, 2007: A multimodel simulation of pan-Arctic hydrology. *J. Geophys. Res.*, **112**(G4), doi:10.1029/2006jg000303.

Smith, M.B., V. Koren, Z.Y. Zhang, Y. Zhang, S.M. Reed, Z.T. Cui, F. Moreda, B.A. Cosgrove, N. Mizukami, E.A. Anderson et al., 2012: Results of the DMIP 2 Oklahoma experiments. *J. Hydrol.*, **418**: 17-48, doi:10.1016/j.jhydrol.2011.08.056.

Sokolov, A.P., D.W. Kicklighter, J.M. Melillo, B.S. Felzer, C.A. Schlosser and T.W. Cronin, 2008: Consequences of considering carbon-nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle. *J. Climate*, **21**(15): 3776-3796, doi:10.1175/2008jcli2038.1.

Sorooshian, S., J.L. Li, K.L. Hsu and X.G. Gao, 2012: Influence of irrigation schemes used in regional climate models on evapotranspiration estimation: Results and comparative studies from California's Central Valley agricultural regions. *J. Geophys. Res.*, **117**, doi:10.1029/2011jd016978.

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**(G1), doi:10.1029/2007jg000562.

Swenson, S. and P.C.D. Milly, 2006: Climate model biases in seasonality of continental water storage revealed by satellite gravimetry. *Water Resour. Res.*, **42**(3), doi:10.1029/2005wr004628.

Swenson, S. and J. Wahr, 2006: Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.*, **33**(8), doi:10.1029/2005gl025285.

Tapley, B.D., S. Bettadpur, J.C. Ries, P.F. Thompson and M.M. Watkins, 2004: GRACE measurements of mass variability in the Earth system. *Science*, **305**(5683): 503-505, doi:10.1126/science.1099192.

Thornton, P.E., S.C. Doney, K. Lindsay, J.K. Moore, N. Mahowald, J.T. Randerson, I. Fung, J.F. Lamarque, J.J. Feddema and Y.H. Lee, 2009: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model. *Biogeosciences*, **6**(10): 2099-2120.

Thornton, P.E., J.F. Lamarque, N.A. Rosenbloom and N.M. Mahowald, 2007: Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. *Global Biogeochem. Cy.*, **21**(4), doi:10.1029/2006gb002868.

Thornton, P.E. and N.A. Rosenbloom, 2005: Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecol. Model.*, **189**(1-2): 25-48, doi:10.1016/j.ecolmodel.2005.04.008.

Timlin, D., M. Kouznetsov, D. Fleisher, S.-H. Kim and V.R. Reddy, 2009: Simulation of nitrogen demand and uptake in potato using a carbon-assimilation approach. In: L. Ma, L.R. Ahuja and T.W. Bruulsema (Editors), *Quantifying and Understanding Plant Nitrogen Uptake for Systems Modeling*. CRC Press, pp. 219-243.

Troy, T.J., E.F. Wood and J. Sheffield, 2008: An efficient calibration method for continental-scale land surface modeling. *Water Resour. Res.*, **44**(9), doi:10.1029/2007wr006513.

U.S. EPA, 2008: Methods for evaluating wetland condition: Nutrient loading. *EPA-822-R-08-023*, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 43 pp.

van den Hurk, B., M. Best, P. Dirmeyer, A. Pitman, J. Polcher and J. Santanello, 2011: Acceleration of land surface model development over a decade of GLASS. *Bull. Amer. Meteor. Soc.*, **92**(12): 1593-1600, doi:10.1175/Bams-D-11-00007.1.

van Dijk, A.I.J.M., L.J. Renzullo and M. Rodell, 2011: Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates

by the Australian water resources assessment system. *Water Resour. Res.*, **47**, doi:10.1029/2011wr010714.

Verdin, K. and S. Jenson, 1996: Development of continental scale DEMs and extraction of hydrographic features, *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM, January 21-26, 1996. National Center for Geographic Information and Analysis, Santa Barbara, CA.

Verstraete, M., 1989: Land surface processes in climate models: Status and prospects. In: A. Berger, S. Schneider and J.C. Duplessy (Editors), *Climate and Geo-Sciences*. NATO ASI Series. Springer Netherlands, pp. 321-340.

Wang, A. and X. Zeng, 2009: Improving the treatment of the vertical snow burial fraction over short vegetation in the NCAR CLM3. *Adv. Atmos. Sci.*, **26**(5): 877-886, doi:10.1007/s00376-009-8098-3.

Wang, Y.P., B.Z. Houlton and C.B. Field, 2007: A model of biogeochemical cycles of carbon, nitrogen, and phosphorus including symbiotic nitrogen fixation and phosphatase production. *Global Biogeochem. Cy.*, **21**(1), doi:10.1029/2006gb002797.

Warner, T.T., 2011: *Numerical weather and climate prediction*. Cambridge University Press, Cambridge; New York, xxii, 526 pp.

Wei, H.L., Y.L. Xia, K.E. Mitchell and M.B. Ek, 2013: Improvement of the Noah land surface model for warm season processes: Evaluation of water and energy flux simulation. *Hydrol. Process.*, **27**(2): 297-303, doi:10.1002/Hyp.9214.

West, T.O., C.C. Brandt, L.M. Baskaran, C.M. Hellwinckel, R. Mueller, C.J. Bernacchi, V. Bandaru, B. Yang, B.S. Wilson, G. Marland et al., 2010: Cropland carbon fluxes in the United States: increasing geospatial resolution of inventory-based carbon accounting. *Ecol. Appl.*, **20**(4): 1074-1086, doi:10.1890/08-2352.1.

Williams, J.R., 1995: The EPIC model. In: V.P. Singh (Editor), *Computer models of watershed hydrology*. Water Resources Publications, Highlands Ranch, CO, pp. 909-1000.

Wood, E.F., D. Lettenmaier, X. Liang, B. Nijssen and S.W. Wetzel, 1997: Hydrological modeling of continental-scale basins. *Annu. Rev. Earth Planet. Sci.*, **25**: 279-&, doi:10.1146/annurev.earth.25.1.279.

Wood, E.F., J.K. Roundy, T.J. Troy, L.P.H. van Beek, M.F.P. Bierkens, E. Blyth, A. de Roo, P. Döll, M. Ek, J. Famiglietti et al., 2011: Hyperresolution global land surface

modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res.*, **47**(5): W05301, doi:10.1029/2010wr010090.

Xia, Y., B.A. Cosgrove, M.B. Ek, J. Sheffield, L. Luo, E.F. Wood, K. Mo and N. team, 2013: Overview of the North American Land Data Assimilation System (NLDAS). In: S. Liang, X. Li and X. Xie (Editors), *Land Surface Observation, Modeling and Data Assimilation*. World Scientific, pp. 337-377.

Xia, Y., C.D. Peter-Lidard, M. Huang, H. Wei and M. Ek, 2014a: Improved NLDAS-2 Noah-simulated hydrometeorological products with an interim run. *Hydrol. Process.*, doi:10.1002/hyp.10190.

Xia, Y., J. Sheffield, M.B. Ek, J.R. Dong, N. Chaney, H. Wei, J. Meng and E.F. Wood, 2014b: Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.*, **512**: 107-125, doi:10.1016/j.jhydrol.2014.02.027.

Xia, Y.L., M. Ek, H.L. Wei and J. Meng, 2012a: Comparative analysis of relationships between NLDAS-2 forcings and model outputs. *Hydrol. Process.*, **26**(3): 467-474, doi:10.1002/Hyp.8240.

Xia, Y.L., K. Mitchell, M. Ek, B. Cosgrove, J. Sheffield, L.F. Luo, C. Alonge, H.L. Wei, J. Meng, B. Livneh et al., 2012b: Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *J. Geophys. Res.*, **117**, doi:10.1029/2011jd016051.

Xia, Y.L., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L.F. Luo, C. Alonge, H.L. Wei, J. Meng et al., 2012c: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**(D3): D03109, doi:10.1029/2011jd016048.

Yang, Z.-L., 2004: Modeling land surface processes in short-term weather and climate studies. In: X. Zhu, X. Li, M. Cai, S. Zhou, Y. Zhu, F.-F. Jin, X. Zou and M. Zhang (Editors), *Observation, Theory and Modeling of Atmospheric Variability*. World Scientific Series on Meteorology of East Asia, World Scientific, New Jersey, pp. 288-313.

Yang, Z.-L. and R.E. Dickinson, 1996: Description of the biosphere-atmosphere transfer scheme (BATS) for the soil moisture workshop and evaluation of its performance. *Global Planet. Change*, **13**(1-4): 117-134, doi:10.1016/0921-8181(95)00041-0.

Yang, Z.-L., R.E. Dickinson, A. Henderson-Sellers and A.J. Pitman, 1995: Preliminary study of spin-up processes in land surface models with the first stage data of Project

for Intercomparison of Land Surface Parameterization Schemes Phase 1(a). *J. Geophys. Res.*, **100**(D8): 16553-16578, doi:10.1029/95jd01076.

Yang, Z.-L. and G.-Y. Niu, 2003: The versatile integrator of surface and atmosphere processes: Part 1. Model description. *Global Planet. Change*, **38**(1-2): 175-189, doi:10.1016/S0921-8181(03)00028-6.

Yang, Z.-L., G.-Y. Niu, K.E. Mitchell, F. Chen, M.B. Ek, M. Barlage, L. Longuevergne, K. Manning, D. Niyogi, M. Tewari et al., 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *J. Geophys. Res.*, **116**, doi:10.1029/2010jd015140.

Young, R.A., C.A. Onstad, D.D. Bosch and W.P. Anderson, 1989: AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. *J. Soil Water Conserv.*, **44**(2): 168-173.

Zampieri, M., E. Serpetzoglou, E.N. Anagnostou, E.I. Nikolopoulos and A. Papadopoulos, 2012: Improving the representation of river-groundwater interactions in land surface modeling at the regional scale: Observational evidence and parameterization applied in the Community Land Model. *J. Hydrol.*, **420**: 72-86, doi:10.1016/j.jhydrol.2011.11.041.

Zeng, X.B. and M. Decker, 2009: Improving the numerical solution of soil moisture-based Richards equation for land models with a deep or shallow water table. *J. Hydrometeorol.*, **10**(1): 308-319, doi:10.1175/2008jhm1011.1.

Vita

Xitian Cai was born in Guangdong, China. He entered Wuhan University in the fall of 2001 where he earned a Bachelor of Engineering degree in agricultural water resources engineering in 2005. He then worked for one year in an engineering company where he designed hydropower plants, reservoirs, and channels. He earned his Master of Engineering degree in hydrology and water resources engineering from Beijing Normal University in 2009. He entered The University of Texas at Austin in August 2009 to work in the fields of climate, hydrology, and biogeochemistry.

Permanent email address: Xitian.Cai@gmail.com

This dissertation was typed by the author.