This is to certify that the
dissertation entitled

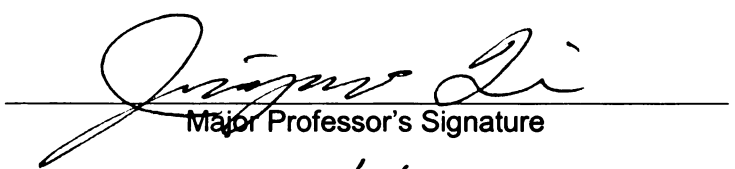
IMPROVING REGIONAL CLIMATE MODELING IN EAST AFRICA USING REMOTE SENSING PRODUCTS

presented by

Jianjun Ge

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Department of Geography



Major Professor's Signature

5/31/2007

Date

IMPROVING REGIONAL CLIMATE MODELING IN EAST AFRICA USING REMOTE SENSING PRODUCTS

By

Jianjun Ge

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Geography

2007

ABSTRACT

IMPROVING REGIONAL CLIMATE MODELING IN EAST AFRICA
USING REMOTE SENSING PRODUCTS

By

Jianjun Ge

Accurate representation of the land surface in regional climate models is becoming unprecedentedly crucial as numerous studies are now focused to simulate the influences of human modification of the Earth's surface on regional and global climate. The unique advantages of remote sensing technique in monitoring the land surface have been recognized for decades. The climate modeling community, however, has yet to fully utilize this technique, especially the recently developed remote sensing products which have been proven to be more suitable for global change studies. The objectives of this study are to improve the representation of the land surface and to investigate the impacts of land cover classification accuracy on regional climate modeling in East Africa.

Several land cover datasets from different sources now exist in almost any region of the world. A new statistical measure Q is developed to evaluate the land cover classification specifically for the purpose of climate modeling. In terms of this Q measure, Global Land Cover 2000 (GLC2000) is ranked the best among four land cover datasets.

To better represent the land surface newly developed MODIS Leaf Area Index (LAI) and Vegetation Fractional Cover (VFC) imageries are incorporated

directly in the Regional Atmospheric Modeling System (RAMS). The default land cover dataset is updated by GLC2000 as well. The impact is examined by comparing the model simulated land surface temperature (LST) and precipitation with MODIS LST and precipitation from the Tropical Rainfall Measuring Mission (TRMM) satellite. This study finds that the incorporation of MODIS LAI and VFC greatly improves the spatial and temporal characteristics of LST. The precipitation, however, is less sensitive to the improved land surface conditions.

The uncertainty originating from the land surface and its propagation need to be examined to truly improve the representation of the land surface in climate models. This study focuses on the land cover classification accuracy, which is the first such investigation. This study finds that classification accuracy under 80% has significant impact on simulated precipitation, especially when the land surface has a greater control of the overlying atmosphere. As the accuracy worsens, the effect becomes much stronger. In remote sensing community, an 85% overall accuracy has been brought up as a guideline of classification quality control. This study shows that this accuracy target can indeed satisfy the requirement of climate modeling in the East Africa region. In reality, however, the classification accuracy can be much lower as historically reconstructed and future projected land cover datasets are extensively used in many climate modeling studies.

To my family

ACKNOWLEDGEMENTS

Thanks to my advisor, Jiaguo Qi, for his good ideas and for helping me through bumps in the road. Thanks to members of my committee, David Lusch, Jeffrey Andresen, Brent Lofgren, and Lijian Yang. My thanks to the co-authors of the papers coming out of this dissertation: Nathan Moore, Nathan Torbick, and Jennifer Olson. Thanks to all the students and project members I have worked with over the years. I appreciate your help and input.

Funding for this study was provided by the National Science Foundation (NSF) Biocomplexity of Coupled Human and Natural Systems Program award BCS 0308420 entitled "An Integrated Analysis of Regional Land-Climate Interactions", and by the NASA's Land Cover and Land Use program through a grant to Michigan State University (NNG05GD49G).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Climate modeling and the importance of land surface

Modern climate change is dominated by human influences, which are now likely large enough to exceed the bounds of natural variability (IPCC, 2001). It is believed that climate changes resulting from human impacts are slow to develop and, therefore, may not become apparent until their effects have become irreversibly advanced. At global scale, anthropogenic emissions of greenhouse gases, such as CO_2 that results from the combustion of fossil fuels, change the atmospheric composition and influence the natural flows of radiant energy. At regional scale, human activities are transforming the land surface at an accelerated pace, which is often more pronounced where people live, work, and grow food. Deforestation in Amazonia (Hahmann and Dickinson, 1997) and desertification in the Sahel (Xue, 1997) are two instances where evidence suggests there is likely to be significant human influence on regional climate.

Over the past four decades, climate models have been employed as primary tools to enhance understanding of past climate changes and to aid prediction of future climates. In climate models, the behavior of the atmosphere is described using a set of differential equations that describe external forcings to the system and the response of the atmosphere to the forcing. Historically, modeling of the atmospheric component of the climate system has received the

most attention. As the availability of computers has become greater and computer storage capacity and speed of computation have increased, climate models have increased in complexity. More efforts have been concentrated on including four major components – atmosphere, oceans, sea ice and land surface – and the interactions and feedbacks among them (McGuffie and Henderson-Sellers, 2001; Washington and Parkingson, 2005).

The land surface is the lower boundary of the atmosphere and thus a key component of climate models. It controls the partitioning of available energy at the surface between sensible and latent heat, and it also controls the partitioning of available water between evaporation and runoff. Lewis F. Richardson was the father of today's climate models. In the first description of a method for constructing a weather forecasts numerical calculations Richardson recognized the importance of the land surface (Richardson, 1922). He noted:

> "The atmosphere and the upper layers of the soil or sea form together a united system.
>
> Leaves, when present, exert a paramount influence on the interchange of moisture and heat. They absorb the sunshine and screen the soil beneath. Being very freely exposed to the air they very rapidly communicate the absorbed energy to the air, either by raising its temperature or by evaporating water into it. A portion of rain, and the greater part of dew, is caught on foliage and evaporated there without ever reaching the soil. Leaves and stems exert a retarding friction on the air."

These is increasing evidence from modeling studies that the influence of the land surface on climate is significant and that changes in the land surface can

influence regional to global scale climate on time scales from days to millennia. Studies have demonstrated the sensitivity of climate to land surface albedo (Charney et al., 1977; Cunnington and Rowntree, 1986; Laval, 1986; Sud and Fennessy, 1982; Lofgren, 1995). Recently, Betts (2000) shows that the positive radiative forcing induced by decreases in albedo can offset the negative forcing that is expected from carbon sequestration by afforestation. Studies have found a change in global temperature and precipitation following a prescribed change in leaf area index (LAI) (Chase et al., 1996). Precipitation patterns in the tropics are altered substantially in a Global Circulation Model (GCM), where two 10-year simulations were performed: one with the current global seasonally varying LAI and one with the potential seasonally varying leaf area index (Nemani et al., 1996). A new area of recognition in climate models is the role of plant roots. Zeng et al. (1998) found, for example, that the root distribution influences the latent heat flux over tropical land. Kleidon and Heimann (2000) concluded that in order to realistically represent the tropical climate system deep-rooted vegetation must be adequately represented.

Instead of focusing on a single parameter change, many studies examined the climate change caused by land use/cover change (LUCC) (Chase et al., 1996, 2000; Pielke et al., 1999; Betts, 2000; Zhao et al., 2001a,b; DeFries et al., 2002; Taylor et al., 2002; Feddema et al., 2005), deforestation (Bonan et al., 1992; Henderson-Sellers et al., 1993; Lean and Rowntree, 1997), and

desertification (Xue, 1997; Nicholson et al., 1998). All these studies have found significant continent-scale changes in temperature, rainfall and other variables.

1.2. The role of remote sensing

Ever since Deardorff's (1978) pioneering work on parameterizations for both soil and vegetation, land surface submodels within climate models have evolved from quite simple treatments of the surface energy, moisture and momentum exchanges to increasingly complex descriptions (Dickinson, 1991; Sellers et al., 1997; Pitman, 2003; Yang, 2003). Also, our ability to sense characteristics of the land surface remotely has improved dramatically, enabling much better data to be input to the more sophisticated parameterizations.

The land surface in climate models is usually represented by a discrete set of land cover types, each characterized by a suite of biophysical parameters (e.g. LAI, vegetation fractional cover (VFC), albedo, root depth and roughness length). Until the last decade, the land cover products used in most climate models were initially compiled from maps, ground surveys, and various national sources, which have inherent limitations (Mathews 1983, Olson et al. 1983, Cihlar 2000). In the mid-1990s, global-scale land cover products generated from remotely sensed images became available. The Global Land Cover Characteristics Database (Loveland et al. 2000), generated from one year 1 km Advanced Very High-Resolution Radiometer (AVHRR) data, have been widely

implemented in various major soil-vegetation-atmosphere transfer (SVAT) schemes, such as the Biosphere-Atmosphere Transfer Scheme (BATS; Dickinson et al. 1986), the Simple Biosphere Model (SiB and SiB2; Sellers et al. 1986, Sellers et al. 1996a, b), the Land Ecosystem-Atmosphere Feedback Model (LEAF; Lee 1992, Walko et al. 2000), and the National Center for Atmospheric Research Land Surface Model (NCAR LSM; Bonan 1996, Oleson and Bonan, 2000)

For more accurate specification of landscape patterns, many climate modeling studies have also attempted to use biophysical variables derived from AVHRR imagery, particularly spatial and seasonal vegetation structure information. For example, AVHRR normalized difference vegetation index (NDVI) was used to determine the temporal variation of LAI and fraction of absorbed photosynthetically active radiation (fPAR) in the revised Simple Biosphere Model (SiB2) (Sellers et al. 1996a,b). Oleson and Bonan (2000) studied the effect of plant functional type and LAI derived from AVHRR imagery on the simulation of surface fluxes for boreal forest using the NCAR Land Surface Model and found a substantial impact of spatial heterogeneity. LAI from AVHRR NDVI was incorporated into the Regional Atmospheric Modeling System (RAMS) to investigate the sensitivity of regional climate simulations to changes in vegetation (Lu 2002).

For many years, terrestrial ecosystem monitoring at moderate spatial resolutions suitable for climate studies, relied almost exclusively on AVHRR imagery. However, reliance on AVHRR imagery with its associated spectral and geometric constraints has limited the ability of the land research community to develop the range of products needed for global change research (Cihlar 1997). Now a new sensor, the Moderate Resolution Imaging Spectroradiometer (MODIS), is providing a series of products of unparalleled quality and sophistication for biophysical observation of the terrestrial environment. MODIS is designed to satisfy the requirements of three different disciplines: atmosphere, ocean and land, with spectral bands and spatial resolution selected to meet different observational needs (Salomonson et al. 1989).

Standard products produced from MODIS imagery include LAI, fraction of photosynthetically active absorbed radiation (fPAR), enhanced vegetation index (EVI), land surface temperature (LST), net primary productivity (NPP), land cover, albedo, etc (Justice et al. 2002). Initial results of validation have suggested that there is good agreement between MODIS LAI product values and field measurements as well as those scaled-up from very high spatial resolution satellite data (Myneni et al. 2002). Through the BigFoot project (Running et al. 1999), the MODIS LAI algorithm has been assessed for distinct biomes over the world, such as savannah and shrub in southern Africa (Tian et al. 2002), broadleaf forest in eastern United States (Shabanov et al. 2003), an agricultural

and broadleaf forest site in North America (Cohen et al. 2003), and the tundra of Alaska (Verbyla 2005).

The enhanced vegetation index (EVI) was developed to optimize the vegetation signal to improve vegetation monitoring by considering the canopy background signal and a reduction in atmospheric influences. The MODIS EVI has been demonstrated to have improved capability to capture multi-temporal vegetation variations, land cover variations, and biophysical parameter variations (Huete et al. 2002).

Along with the new MODIS land cover products (Friedl et al. 2002), the Global Land Cover 2000 dataset (GLC2000) has recently become available (Fritz et al. 2003). Sponsored by the Joint Research Center (JRC) of the European Commission (EC), GLC2000 was developed from SPOT VEGETATION 1 km data with enhanced spectral, spatial, radiometric, and geometric quality. Compared to the AVHRR-based land cover datasets, these two products are more suitable for monitoring land surface properties at regional to global scales and have great potential to be employed in climate modeling systems in the future (Latifovic et al. 2004, Giri et al. 2005).

Remote sensing data can not only be used to provide better initial and boundary conditions, but also to evaluate the model performance and the accuracy of the forecasts. Currently, validation of global and regional climate

models is based on the comparison between model outputs of standard meteorological fields and meteorological observations. The main problem with traditional meteorological observations when used to validate models is their poor representation of the grid-point average simulated by a model. Well documented station observations of climate variables are mainly located in populated and industrialized regions. Observations are very sparse both spatially and temporally over many rural regions of the world, which brings inevitable challenges for climate modeling studies.

Several climatic variables can now be measured from space. Rainfall and land surface temperature (LST) are of the two most important ones. Since late 1997, the Tropical Rainfall Measuring Mission (TRMM; Kummerow et al. 2000) has been successfully collecting passive microwave and radar-derived tropical precipitation data over both the land and ocean. The latest rainfall data product (3B42 version 6) has a quarter-degree spatial resolution and 3-hour temporal resolution covering from 50 degree south to 50 degree north globally. Although the primary objective of TRMM was to improve climate models and to aid them in climate prediction (Kummerow et al. 2000), very few studies have fully utilized this dataset for model validation and rainfall data assimilation.

AVHRR derived LST was compared with the output from the NCAR Community Climate Model version 2 (CCM2) coupled with BATS to illustrate the differences between air temperature and skin temperature at the global scale (Jin

et al. 1997). Since then, few other efforts have attempted to apply satellite LST in climate modeling. This is partly because the algorithm of AVHRR LST was a simple extension of the sea surface temperature (SST) method, which sometimes produced to unacceptable errors (Price 1984, Becker, 1987). AVHRR data are not suitable for accurate cloud detection over land which is required for an operational LST algorithm. With multiple thermal-infrared bands specifically designed for LST retrievals. MODIS is providing a much more accurate LST product for both regional and global climate modeling (Wan et al. 2002, 2004).

1.3. Land cover accuracy and uncertainty assessment

Modeling of the climate system which has a wide variety of components is a formidable task, and as a result climate models have uncertainties (Mahlman 1997, Palmer 2000, Vidale et al. 2003). Simulated climatic variables differ considerably between models. The uncertainties in the climate models originate from several components, of which the SVAT scheme is believed to be very important. The primary aim of SVAT is to take into account the essential land surface processes and provide estimates of latent and sensible heat fluxes between the land surface and atmosphere. The uncertainties in SVAT schemes arise from three elements: the model parameterization, meteorological forcing data, and vegetation/soil inputs.

In SVAT schemes, land surface features are differentiated by land cover types, and then biophysical parameter sets are assigned to each land cover type. While numerous studies have been conducted to investigate the sensitivity of the climate model to some particular biophysical parameters (e.g., Franks et al. 1997, Bastidas et al. 1999, Lynch et al. 2001), some attention has been given to the uncertainty stemming from the land cover dataset. By examining the impact of different land cover datasets (digitalized vs. satellite derived) on the predicted variables, Molders et al. (1997) found that the distribution of daily averages of temperature and humidity changed less than 0.2 °C (1%) and 0.2 g/kg (1.5%) if the coverage of the various land-use and soil types differed by only about 5% between the datasets. Pauwels and Wood (2000) compared the effect of differences in spatial resolution of land cover data to land-atmosphere model results relative to the effect of differences in land cover sensors and classification schemes. They found that the uncertainty in model results arises mainly from the land cover classification and that the lack of spatial resolution had a lower effect. Overall, an uncertainty of approximately 15% in modeled energy and water fluxes and states has to be assigned in evaluating the model simulation.

Until all important biophysical parameters in any SVAT scheme can be mapped directly by satellites, land cover datasets will continue to play a key role in representing the land surface conditions. No land cover product is 100% accurate even those produced from the most advanced remote sensing imagery. The uncertainties caused by the classification inaccuracy thus need to be

investigated. Because several land cover datasets are sometimes available for a study, the effect of land cover classification accuracy on model results can help climate modelers to decide which one to select. For land cover dataset producers, the effect of classification accuracy can be used as a guideline for the accuracy assessment and quality control. If the effect is not significant, no extra efforts are needed to pursue higher mapping accuracies.

In the remote sensing community, some target accuracy thresholds have recently been recommended in an attempt to provide guidelines to the classification quality. Thomlinson et al. (1999), for example, set as a target an overall accuracy of 85% with no class less than 70% accurate. However, classification accuracy is usually interpreted differently from the viewpoint of various users. The effect of land cover accuracy for a particular application, such as climate modeling in this study, remains an unanswered question. The accuracy targets commonly specified have largely not been tested from the perspective of the operational use of land cover data. As anthropogenic impacts on the land surface and climate at various scales are attracting more and more attention from both scientists and policy makers, an effective way to evaluate classification accuracy needs to be developed specifically for climate modeling studies and the impact of classification uncertainty on climate simulations needs to be fully investigated.

1.4. Research objectives

The overall research objectives in this study are to improve the representation of the land surface in a regional climate model using newly developed remote sensing products and to address the uncertainty of land cover classification and its impact on model simulations. The specific research objectives are:

1. To develop a new method to evaluate various land cover products for regional climate modeling
2. To improve the representation of the land surface in a regional climate model using the highest-quality land cover product from Objective 1 and newly developed MODIS biophysical products (LAI and VFC).
3. To investigate the land cover classification uncertainty and its impact on regional climate simulations.

The new classification accuracy evaluation method is designed specifically for regional climate modeling, addressing the shortcomings of traditional methods used in the remote sensing community. Impacts of improved land surface representation will be examined by comparing model outputs with remote sensing observations (TRMM and MODIS LST). This will be the first study to directly incorporate MODIS biophysical products into a regional model and to use TRMM and MODIS LST to evaluate the model performance in a region where station observations are extremely scarce. Also, a classification accuracy

threshold will be identified by the third objective. This will provide an additional guideline for land cover accuracy assessment in the remote sensing community. When the classification accuracy is low, climate simulations must be interpreted with caution, especially when historically-developed and future-projected land cover products are used.

The dissertation is outlined as follows:

Chapter 2 describes the methodology of this study. A more detailed discussion of the various remote sensing products is provided and where and how they are used is illustrated. In addition, the regional climate model used in this study is briefly introduced.

Chapter 3 (Objective 1) focuses on the theoretical development and application of a new statistical method to evaluate land cover classification accuracy. The most appropriate land cover product is identified from several existing ones.

Chapter 4 (Objective 2) describes the implementation of the new land cover product identified by Objective 1 and new MODIS LAI and VFC products. Spatial and temporal characteristics of the improved land surface map are compared with the default representation in the regional model. The impact of the

improved land surface characterization is examined by comparing model-simulated LST and precipitation with MODIS LST and TRMM data respectively.

Chapter 5 (Objective 3) investigates the impacts of land cover classification accuracy on simulated precipitation. A threshold of accuracy is identified.

Chapter 6 is the conclusion. Future studies in this area are discussed.

Chapter 2

Research Design

2.1. Regional Atmospheric Modeling System (RAMS) overview

The regional climate model used for the numerical simulations in this study is the Regional Atmospheric Modeling System (RAMS) Version 4.4 (Pielke et al. 1992, Cotton et al. 2003). RAMS is a three-dimensional, nonhydrostatic, general purpose atmospheric simulation modeling system, which solves equations of motion, heat, moisture, and mass continuity in a terrain-following coordinate system. It is an atmospheric model which is capable of both numerical weather prediction and regional climate simulation.

The SVAT scheme employed in RAMS is the Land Ecosystem-Atmosphere Feedback model, version 2 (LEAF-2) (Lee 1992, Walko et al. 2000). LEAF-2 represents the storage and vertical exchange of water and energy in multiple soil layers, temporary surface water or snow cover, and vegetation and canopy air. The special feature of LEAF-2 is its ability to represent fine-scale surface variations by dividing surface grid cells into sub-grid patches, which are assigned based on the land cover types in a model grid cell. Each patch has one land cover type and influences the overlying atmosphere in its own unique way according to its fractional area of coverage. The biophysical characteristics, such as albedo, leaf area index, fractional vegetation cover, etc., are prescribed for the land cover type in each patch. However, as in many other SVAT schemes

seasonal variation of vegetation is characterized by simple mathematic equations.

2.2. Research design

The following flowchart summarizes the research design in this study.

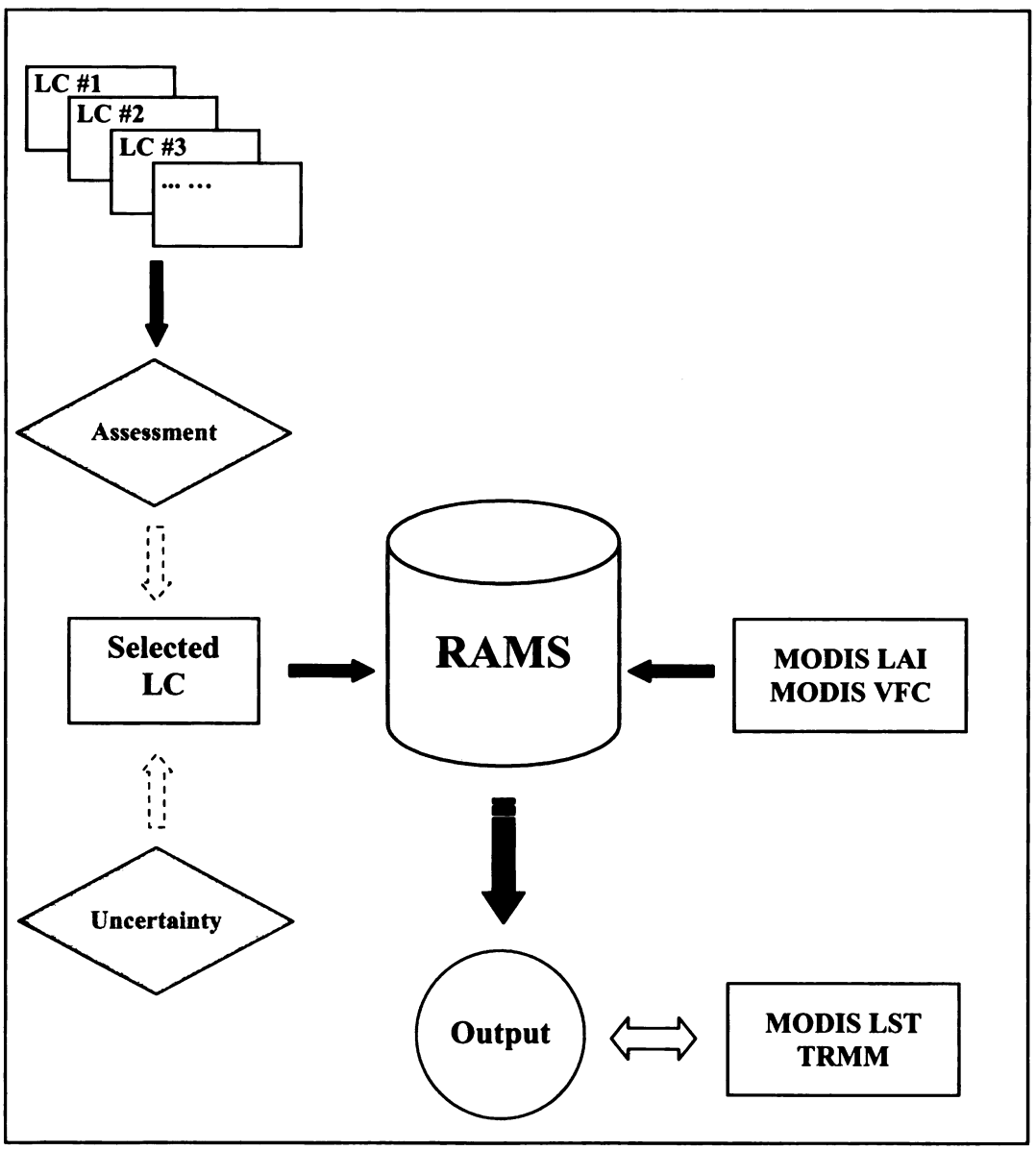


Figure 2-1. Flowchart of research design in this study

A new land cover product and one-year monthly MODIS LAI & VFC datasets will be directly ingested into RAMS to improve the land surface conditions both spatially and temporally, and the effects of these new land cover and biophysical products will be examined separately. RAMS will be run three times: As a control, the first run uses the default land cover dataset and prescribed LAI & VFC. The second run uses a new land cover product, but prescribed LAI & VFC. For the third run, both new datasets will be used. Model-simulated precipitation and LST will be compared with TRMM precipitation and MODIS LST. RAMS does not have LST as an output variable. An algorithm thus will be used to calculate LST based on the temperature of vegetation and soil in a grid cell.

A new statistical evaluation method is developed and applied to several existing land cover products to select the most appropriate one for climate modeling in the study region. Therefore, various land cover products will be evaluated and compared before running the RAMS model.

To study the uncertainty effect of land cover classification inaccuracy, a range of misclassifications (5% - 50% at 5% interval) will be simulated on the baseline land cover dataset. During this process, spatial locations (excluding water) will be randomly selected and original land cover types will be converted to random predominant types. RAMS will then run multiple times, each with a

different amount of classification error. Model outputs will be compared and the threshold of classification accuracy will be identified. Also, the influence of model configuration will be examined.

Images in this dissertation are presented in color.

Chapter 3

Biophysical Evaluation of Four Land Cover Products for Regional Climate Modeling

3.1 Introduction

3.1.1 Study area

This study focuses on the East Africa region (Figure 3-1), which covers Kenya, Uganda, Rwanda and Burundi, as well as parts of Congo and Tanzania and a small section of the Indian Ocean. The geographic coordinates for the upper-left and lower-right corners of this study area are (3.71°, 23.80°) and (-9.34°, 41.80°), and the size is about 2000km by 1500km. This region contains some of the most varied topography in the world, including large lakes (e.g., Lake Victoria), rift valleys, and snow-capped mountains (e.g., Mount Kilimanjaro) on the equator. This region also has a wide range of land use/cover types: from savannas to forest to intense agriculture. Rainfall in East Africa mostly occurs during the boreal spring (long rains, March-May) and autumn (short-rains, September/October-December) seasons as the intertropical convergence zone (ITCZ) migrates through the equator from south to north, and vice versa.

The land surface in this region has been substantially changed by human activities during the past decades due to increased population and other factors, which may likely trigger climate change. At the same time, people in this region

are highly vulnerable to climate variability and change. Much of the region is arid, semi-arid or sub-humid and yet highly populated by pastoralists and agro-pastoralists dependent on rainfed pasture and crops. This study is in the context of the Climate-Land Interaction Project (CLIP), which integrates diverse methods to quantify the two-way interactions between land use and regional climate systems.

3.1.2 Land cover classification evaluation

To represent the land surface conditions realistically, RAMS needs an accurate land cover dataset as input. Traditional land cover classification accuracy assessment is primarily based on ground-based surveys or interpretation of high spatial resolution aerial photos and satellite images. By comparing the classified land cover with ground-truth data, error metrics can be developed to report the commission and omission errors (Congalton 1991). Shortcomings of this traditional method are apparent. Firstly, it is not cost effective. Secondly, reported accuracy is highly susceptible to sampling error when the sample taken is far sufficient to represent the study area, especially when classifications have regional and global coverage. Thirdly, the ground-truth data are usually obtained at one particular time, which can't take into account the phenological consistency within each class. Finally, classification accuracy is often interpreted differently from different users (Lark 1995, Brown et al. 1999). The traditional method is not designed for the purpose of climate modeling.

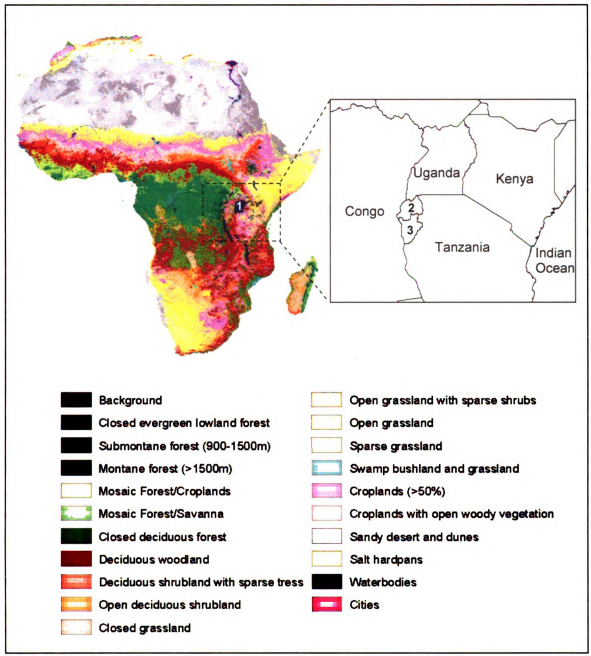


Figure 3-1. Study area illustrated by GLC2000 for Africa. Area 1, 2, and 3 refer to Lake Victoria, Rwanda, and Burundi, respectively.

Researchers have proposed to use biophysical products with observable, unambiguous, and continuous structural variables to make an optimized classification for specific needs (Running et al. 1995, DeFries et al. 1997, Cihlar 2000). This idea has not yet been used in the phase of classification assessment.

In this study, a new statistical method, called Q, is developed to biophysically evaluate the seasonal consistency of classification. This method eliminates sampling error and is designed specifically for climate modeling.

3.2 Methodology

In the RAMS model, land cover types in a grid cell are treated as subgrid 'patches', each with its own vegetation, soil, canopy air, and etc. (Walko et al. 2000). The surface characteristics of a patch are then represented by a group of biophysical parameters: LAI, VFC, albedo, etc.

Biophysical parameters are now available from remote sensing products. For each patch, biophysical parameters can be taken directly from corresponding remote sensing products. As a result, variation of these parameters in a patch can be used to indicate if this patch (land cover type) has been appropriately assigned (classified). Aggregating all patch-level variations spatially can be used to evaluate land cover products from the perspective of climate modeling. The

statistical measure Q presented in this paper is developed on LAI which is one of the most important biophysical parameters.

3.2.1 Q development

For an area of interest (a RAMS grid cell) with *k* classes, LAI variation, *V*, within one class *i* can be calculated using statistical error sum of squares (Stapleton 1995).

$$V_i = \sum_{n=1}^{p_i} \left(LAI_{in} - \overline{LAI_i} \right)^2 \qquad (3\text{-}1)$$

where p_i is the total number of pixels for class *i*, *n* refers to any particular pixel, and $\overline{LAI_i}$ is the mean LAI value. Also, seasonal changes of LAI must be taken into account. Including the temporal dynamics of vegetation is expected to improve the evaluation. By adding this temporal LAI information, equation (3-1) can be rewritten as:

$$W_i = \left(\sum_{t=1}^{T} V_{it} \right) / T \qquad (3\text{-}2)$$

where *t* represents time. In this paper two years of monthly LAI data (January 2002 to December 2003) are used with *t* ranging from 1 to 24, and *T* therefore equals the maximum number of time periods used (For some areas not all 24 time periods can be used because of quality control for LAI pixels. This is discussed later.) Summing W_i for all classes will give a good indication of within-

class LAI variation. However, there is still another very important issue left out. For a given geographic area, different land cover products may classify an unequal number of classes to it because of different classification schemes. Therefore, summation of W_i must be normalized by $(N - k)$ to yield the final statistical measure Q. It is expressed as:

$$Q = \frac{1}{(N-k)} \sum_{i=1}^{k} W_i \qquad (3\text{-}3)$$

where N is the total number of pixels for the area of interest (e.g., a grid cell in RAMS) and is the same for different land covers, and k is the total number of classes. It is assumed here that N is larger than one. In most regional climate modeling studies, the spatial resolution is usually coarser than 10km because of the limitation on computing resources. By combining equation (3-1) and (3-2), equation (3-3) can be written as:

$$Q = \frac{1}{(N-k)} \sum_{i=1}^{k} \left(\left(\sum_{t=1}^{T} \sum_{n=1}^{p_i} \left(LAI_{itn} - \overline{LAI}_{it} \right)^2 \right) / T \right) \qquad (3\text{-}4)$$

In summary, Q aggregates LAI variation in each land cover type for each time period for an area. A smaller Q value indicates more consistent biophysical characteristics over a two-year time period and therefore, a more appropriate classification for climate modeling.

3.2.2 Illustration

As an illustration, Figure 3-2 represents a simulated area with 4×4 pixels. The LAI value in each pixel is also presented at one particular time. For the purpose of clarity, in this study this type of rectangular area will be referred as '*quadrate*' to differentiate from '*pixel*' in image processing and '*grid cell*' in climate modeling. One quadrate is composed of smaller image pixels and corresponds to a grid cell. In this illustration two quadrates in Figure 3-2 represent two different classifications for one area. Quadrate a) has two classes and its Q value is 0.02. In comparison, quadrate b) has three classes and its Q value is 0.47. These two Q values suggest quadrate a) is more properly classified.

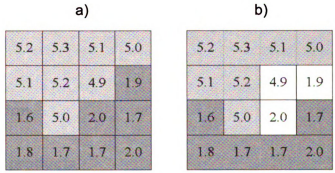


Figure 3-2. A simulated 4×4 pixel area. Quadrate a) has 2 classes and quadrate b) has 3 classes. Numbers shown are LAI values.

3.2.3 Evaluation design

Non-overlapping equal size quadrates are tiled completely for the study area. Q values were calculated quadrate by quadrate for each land cover products. The mean Q value was utilized as the final evaluator. By doing so, the

effect of spatial resolution of the climate model can be studied by adjusting the quadrate size. As quadrate size becomes smaller, the Q value may decrease. In the case where the size of the quadrates is one pixel, there would be just one LAI value and thus no variation at all.

For the climate modeling in this study, the grid spacing of RAMS is 50 km. This resolution was designed based on the topographical characteristics in the study area and computational requirement. The evaluation was extended to quadrate size of 30×30km and 100×100km to study how resolution affects the application of Q. In the following sections, Q values for all four land covers for quadrate size of 30×30km, 50×50km and 100×100km will be calculated and analyzed.

3.3 Data description

3.3.1 Land cover products

Q was applied to the following four land cover products, which exist for climate modeling in East Africa:

1) GLC2000 for Africa (See Figure 3-1)

This product was developed by the Joint Research Centre's Global Vegetation Unit based primarily on SPOT VEGETATION daily 1km data, which

were acquired from 1 November 1999 to 31 December 2000 (Mayaux et al. 2004). Other data sources such as NDVI, radar and Defense Meteorological Satellite Program (DMSP) were also used. GLC2000 uses the Land Cover Classification System (LCCS) developed by the Food and Agriculture Organization of the United Nations (FAO) and the United Nations Environment Programme (UNEP), which contains 27 land cover classes (Di Gregorio et al. 2000). The GLC2000 for Africa was downloaded at http://www-gvm.jrc.it/glc2000. It was in geographic coordinates and then reprojected to Lambert Azimuthal Equal-area projection.

2) MODIS land cover

This product (MOD12Q1) was developed by MODIS Land Cover group at Boston University. It was prepared using MODIS Terra daily data acquired from 15 October 2000 to 15 October 2001, except for June 2001, which is missing due to instrument down time. This product was based on the International Geosphere-Biosphere Programme (IGBP) global vegetation classification scheme which has 17 classes (Friedl et al. 2002). MOD12Q1 for Africa was downloaded from http://duckwater.bu.edu/lc/mod12q1.html and was already in Lambert Azimuthal Equal-area projection with 1km resolution.

3) OGE

This is the Global Land Cover Characterization database (version 2.0) using the Olson Global Ecosystem legend (Olson 1994a, b). It was developed by the U.S. Geological Survey (USGS), the University of Nebraska-Lincoln (UNL), and the European Commission's Joint Research Centre (JRC), based on 1km AVHRR data spanning April 1992 through March 1993. It is archived at the EROS Data Center http://edcdaac.usgs.gov/glcc/glcc.html. This dataset is called OGE for short in this paper.

4) LEAF

This is the default land cover used in LEAF-2 within the RAMS model, termed LEAF herein. LEAF aggregates the OGE dataset into 31 classes to use the BATS land-surface parameters (Walko et al. 2000).

3.3.2 MODIS LAI product

The purpose of this evaluation is to select an appropriate land cover for climate simulations at a time period of interest (2002-2003) rather than to examine whether these land covers were accurately mapped when they were produced. 1km monthly MODIS LAI data from January 2002 to December 2003 were used in this evaluation. The newly reprocessed MOD15_BU LAI C4.1 data were downloaded at ftp://primavera.bu.edu/pub/datasets/MODIS/. They were reprojected to Lambert Azimuthal Equal-area projection with 1km resolution. For LAI pixels with normal values (0-7), Quality Assessment (QA) flags were used to

detect cloud contamination in each pixel (Myneni et al. 2002). Only high quality LAI pixels having QA values from one to four were selected for Q assessment. These pixels were produced by the main algorithm, without saturation under clear sky conditions. For LAI pixels with fill values (200, pixels outside projection; 253, Barren, desert, or very sparsely vegetated; 254, water; 255, non-computed pixels or missing pixels), only 253 and 254 were used by replacing them with zero. It needs to be noted that there is persistent cloud cover over part of the Congo forest in the study area, especially during the wet season. But, this is not a problem since Q can be calculated with only one month of data.

3.4. Results

3.4.1 Q values

Q values were calculated quadrate by quadrate for all four land cover products at three quadrate sizes (30×30, 50×50 and 100×100km). Only Q values at quadrate size of 30×30km are graphically presented here (Figure 3-3). It is noticeable that Q has a similar spatial pattern for all land covers. It has larger values for areas with complex and heterogeneous landscapes, such as ecotone boundaries, and has smaller values for more homogeneous areas, such as forests and deserts. Large water areas like Lake Victoria have Q of zero because of no LAI variation.

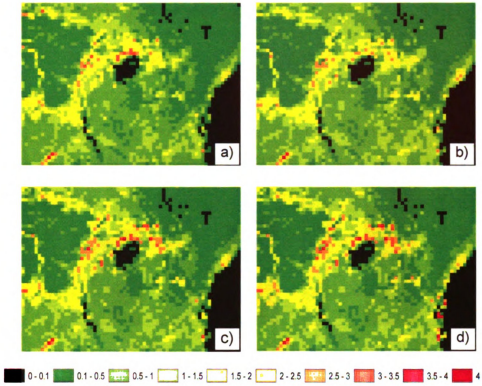


Figure 3-3. Q maps for four land covers at the quadrate size of 30×30 km: GLC2000 (a), MODIS land cover (b), OGE (c), and LEAF (d).

Having Q values at quadrate level, the mean Q value was calculated and used as the indicator of evaluation. The mean Q values for each land cover product at three different quadrate sizes are plotted in Figure 3-4. The mean Q value for GLC2000 is the smallest, while the Q value for LEAF is greater than the other three land covers for every quadrate size. The rank for mean Q values for all four landcovers is $Q_{GLC} < Q_{MODIS} < Q_{OGE} < Q_{LEAF}$. This suggests that GLC2000 has the best land cover classification and LEAF has the least appropriate classification in terms of LAI variation. Also, MODIS land cover is better than OGE. As expected, for a given land cover, Q values increase as the size of

quadrates increases. For example, the mean Q value for GLC2000 increases from 0.69 for 30×30 to 0.80 for 100×100.

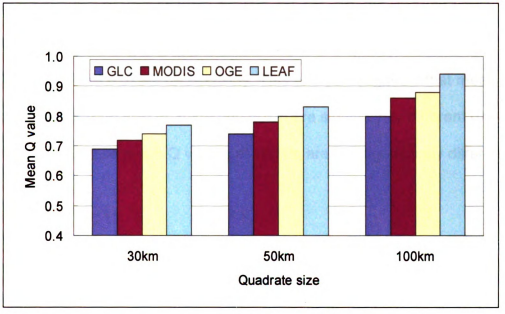


Figure 3-4. Mean Q for all land covers at three scales

3.4.2 Significance test

A null hypothesis that mean Q values from the four land cover products are all equal was tested using the One Way Analysis of Variance (ANOVA) (Scheffe 1959). One advantage of using ANOVA, rather than multiple t-tests, is that it gives one *p* value for a large number of groups. It would have required six pairs for the t-test to evaluate the four land cover classifications in this study. Like other statistical tests, ANOVA assumes that within each sample the values are independent and normally distributed. This may not be satisfied for Q values since they may be spatially correlated, which is common for most spatially

distributed data. In this study, I assumed there was no spatial correlation for Q values. A significance level of 0.05 was chosen to conduct the test for each quadrate size. Sample sizes (number of quadrates) for four land cover classifications are equal on all three scales: 3484 for 30×30km, 1240 for 50×50km and 300 for 100×100km. The resulting p-values for each test are 0.0002, 0.0054, and 0.0739, respectively. This analysis suggests that mean Q values at the quadrate size of 30km and 50km are significantly different (0.0002, 0.0054 << 0.05), while mean Q values at 100km are not significantly different.

In order to identify specific differences between pairs of groups, Tukey's method of the multiple comparison procedure (MCP) was used following the ANOVA (Zhuang et al. 1995, Stapleton 1995). ANOVA only tells whether there is a difference among the four land cover products, but not which ones are significantly different. MCP provides which land cover classification is different if significant difference has been found by ANOVA.

Figure 3-5 presents part of the results from the MCP for quadrate sizes 30×30km (Figure 3-5 a) and 50×50km (Figure 3-5 b). In this figure, the positions of the dots represent the differences between sample means for Q. Parenthesis and dashed lines indicate the extent of the confidence intervals for differences between population means. If a confidence interval does not contain zero, the difference for that pair is significant. As Figure 3-5 illustrates, GLC2000 has a significantly smaller Q value than LEAF at both the 30km and 50km quadrate

sizes, and at 30km quadrate size GLC2000 is also significantly different from OGE. All other paired differences are not significant. Since ANOVA determined that there is no significant difference among the four land cover classifications at the 100×100km quadrate size, it was not necessary to conduct MCP at this scale.

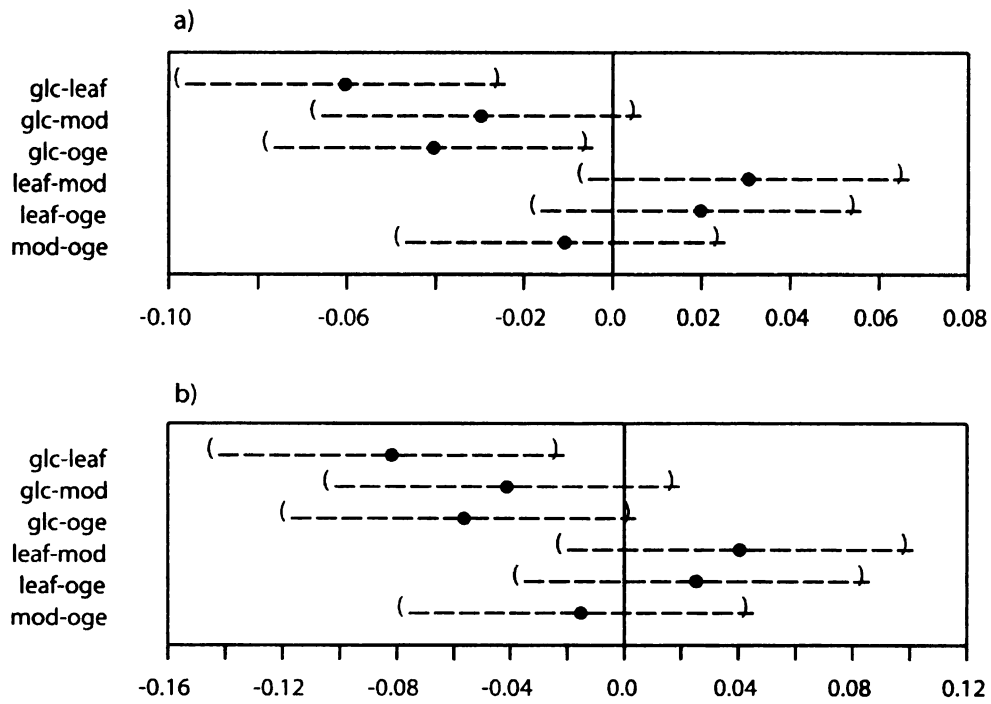


Figure 3-5. Multiple comparison (Tukey's method) results for mean Q values for a quadrate size of 30×30 km (a) and 50×50 km (b). MOD means MODIS land cover. Confidence intervals were built on significance level of 0.05.

3.4.3 Single class investigation

For illustrative purposes, Q was also applied to a single class in GLC2000 to investigate the LAI variation within that class. The croplands (>50%) class was

selected for this study because it is one of the most prevalent land cover types, occupying about 19% of the study area (Figure 3-1). It is defined as regions of intensive cultivation and/or sown pasture.

In Figure 3-6a, Q values of this single class calculated at the quadrate size of 30×30km are presented. Map cells in white represent those quadrates not containing any croplands. The mean Q value for this class is 1.017, which is higher than that for all other classes of GLC2000. This is because agricultural fields in Africa are usually small and mixed with savanna and fallow patches, which preclude a reliable mapping at 1km spatial resolution. The spatial pattern is very similar to the map in Figure 3-3a. High Q values (yellow and red) tend to occur in areas with complex landscapes.

A hotspot with high Q values was identified visually in Figure 3-6a. Geographically, it is in the Mount Elgon area that straddles the border between Kenya and Uganda (Figure 3-1). In the Q map it consists of four contiguous quadrates (total size of 60×60km). The mean Q value for this area is 3.265, which is much higher than the total mean value (1.017). The spatial distribution of the croplands (>50%) class in this hotspot area is presented in Figure 3-6b. It occupies 1069 1km pixels (about 29.7%) in the original GLC2000 dataset. To further investigate the cause of the high Q, Africover land cover at 100m resolution was examined for this patch classified as croplands by GLC2000 (Figure 3-6c). Africover was produced from a visual interpretation of TM data

from the year 2000 to 2001 (Jansen et al. 2003). It is considered as ground truth here because of its much higher resolution (30 x 30m). Africover shows the GLC2000 cropland patch in this hotspot area is actually composed of up to five diversified land cover types: forest, open shrubland, closed shrubland, savanna and crop. Integrating such a complex landscape to one single type will certainly produce high within-class LAI variation.

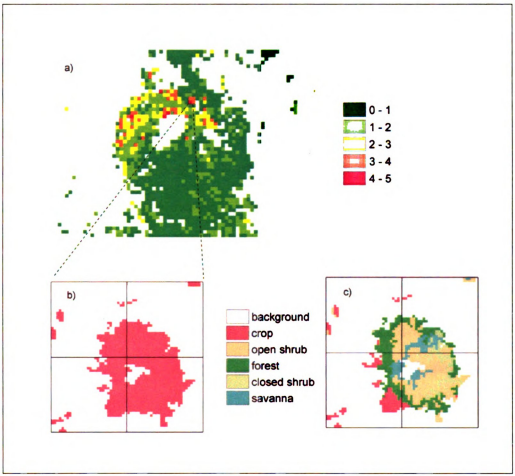


Figure 3-6. Q was applied to a single class, croplands (>50%), in GLC2000. (a) is Q map at 30×30km quadrate size; (b) croplands in the hotspot in (a), pointed by two lines; (c) Africover corresponding to (b).

3.5. Conclusion and discussion

Based on the needs of climate modeling, a statistical measure, Q, was developed to biophysically evaluate four land cover products: GLC2000, MODIS land cover, OGE and LEAF, using the monthly MODIS LAI product. Evaluations were conducted at three spatial scales (quadrate sizes): 30×30, 50×50 and 100×100km. This evaluation found that in terms of Q, GLC2000 ranks the best and LEAF ranks the lowest at every scale. MODIS land cover is better than OGE. As quadrate size increases, the differences between the land cover products tended to decrease. For the quadrate sizes of 30km and 50km, GLC2000 is significantly better (i.e., the smallest mean Q value) than LEAF, and for the quadrate size of 30km GLC2000 is also significantly better than OGE. This suggests that the LEAF dataset (built into the RAMS model) needs to be updated by GLC2000 in order for the model to better capture the surface conditions in East Africa.

There are several comments to be made on the proposed approach. First, the time period of LAI data should be fairly long. Based on the quality control flags, only high quality LAI pixels should be picked. Thus, not every LAI pixel is valid for Q calculation. If for example only a couple of months of LAI are used, there might not be enough high quality LAI pixels to calculate Q in some quadrates. This is especially true for tropical areas, where persistent cloud cover could exist for months.

The second comment is related to the overall quality of the MODIS LAI data. The accuracy of using Q to evaluate alternative land cover classifications depends on the accuracy of LAI product used. Two measures were taken to address this issue. One was using monthly LAI data, composited from 8-day data by selecting high quality pixels over a month period. The accuracy of the 8-day LAI product is about 0.5 LAI (Wang et al. 2004). The other was further filtering the LAI pixels finally used by Q according to the quality assessment flags. Only pixels produced by the main algorithm under clear sky conditions were selected. Nonetheless, the overall LAI quality in this region is still not well known. Some publications have already shown that the LAI product has limitations (Wang et al. 2004).

Finally, LAI is only one of the biophysical variables representing surface properties of land cover in climate modeling. Other biophysical variables (e.g., fractional vegetation cover or albedo) are equally important. Q may also be applied to these variables. More complete conclusions may be drawn by evaluating alternative land cover classifications using a suite of biophysical variables.

Chapter 4

Improving the Land Surface Representation in RAMS and Impacts Analysis

4.1 Introduction

The land surface plays a prominent role in climate modeling, since it exchanges momentum, energy, water, and other important chemical constituents with the atmosphere. The land surface is characterized by pronounced spatial heterogeneity that spans a wide range of scales. In the last decade or so, the importance of representing land surface heterogeneity representation, especially at the subgrid scale, has been increasingly recognized in a large number of climate modeling studies. Ideally, surface heterogeneity effects could be accounted for by running a coupled atmosphere-surface model at a very high spatial resolution, so that the heterogeneity is explicitly captured. However, this approach is too computationally demanding even for most regional modeling. The land surface heterogeneity, therefore, needs to be parameterized within the framework of complex land surface process schemes.

The Land Ecosystem-Atmosphere Feedback model version 2 (LEAF-2), which is the SVAT scheme in RAMS, represents subgrid heterogeneity through a discrete number of homogeneous subregions, also referred to as "patches" or "tiles" (Avissar and Pielke 1989). Patches are selected on the basis of the land cover types (evergreen broadleaf forest, savanna, water, etc), and they directly

exchange fluxes with the atmosphere independently of each other. Each patch occupies a fractional horizontal area of a grid cell in RAMS. Net momentum, moisture, sensible heat, longwave, and shortwave fluxes are integrated over all patches, weighted by the corresponding patch fractional areas. This approach does not keep track of the actual geographical location of the patch within the model grid cell; all the subareas belonging to a given land cover class are treated equally.

The LEAF-2 components and the flux pathways between them are illustrated in Figure 4-1. This example includes two patches beneath an atmospheric column (A) within a single RAMS grid cell. Both patch 1 and patch 2 have partial vegetation cover (V) and patch 2 alone has snowcover (S). Two soil layers (G) and canopy air (C) are also shown in this figure. Fluxed are denoted by smaller letters. The prefix (*w*, *h*, or *r*) indicates whether the flux involves the transfer water, heat or longwave radiation, and the two suffix letters denote the source and the receptor (*g* for ground, *s* for snow, *v* for vegetation, *c* for canopy, and *a* for reference height air). The one exception is wgvc, which denotes a flux of water from the soil to the canopy air through the vegetation by the means of transpiration.

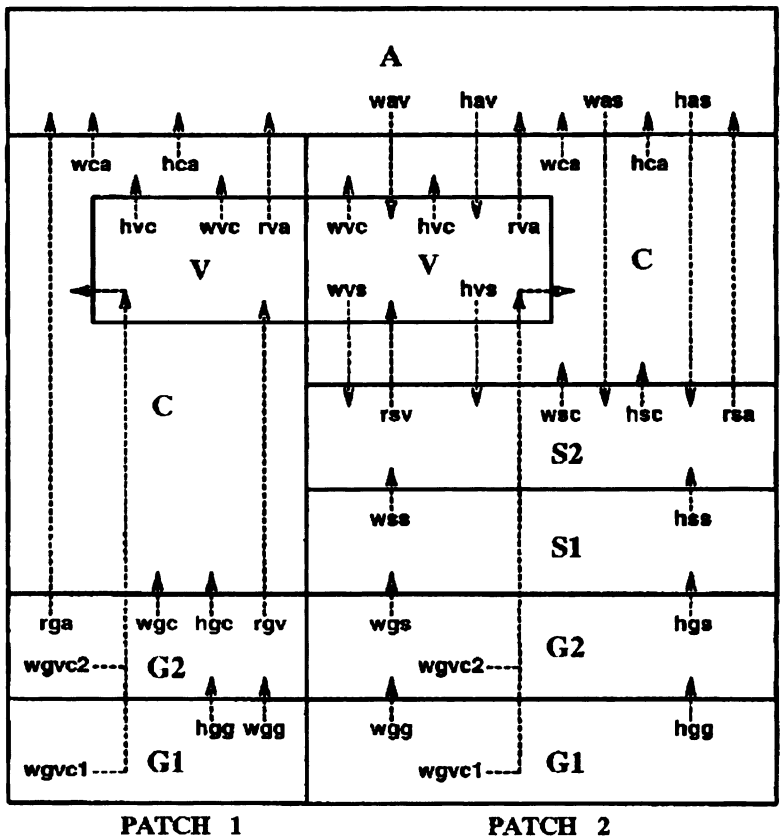


Figure 4-1. Schematic representation of heat and moisture transfer between components of LEAF-2 for two patches within a model grid cell (Walko et al. 2000). See text for detailed description.

In LEAF-2, the standard land cover map that generates the subgrid patches during the initialization of RAMS is the Global Ecosystem dataset (OGE, Olson 1994a, b). Currently, the 94 ecosystem classes in OGE have been cross referenced to the 18 BATS classes (Dickinson et al. 1986) plus some additional classes, which were referred to as LEAF-2 cover in Chapter 1. The RAMS user decides the number of patches, and LEAF-2 fills those patches with the most predominant classes.

These various land cover types (patches) are characterized by a suite of biophysical parameters: leaf area index (LAI), vegetation fractional cover (VFC), displacement height, roughness height, albedo, and emissivity. Some of these biophysical variables are specified from observations in field experiments, while others are educated guesses.

Table 4-1. Biophysical parameters for some important land cover types

LC types	Evergreen Broadleaf Forest	Crop/Mixed Farming	Open Shrubland	Grassland	Woodland
Albedo	0.06	0.20	0.12	0.11	0.08
Emissivity	0.95	0.95	0.97	0.96	0.96
LAI	6.00	6.00	6.00	2.60	5.70
D LAI	1.00	5.50	5.40	2.00	2.30
VFC	0.80	0.85	0.22	0.73	0.80
D VFC	0.10	0.60	0.12	0.11	0.17
Roughness length	2.21	0.06	0.08	0.04	0.83
Displacement height	20.7	0.7	0.2	0.2	7.4
Root depth	1.20	1.00	0.60	0.70	1.00

Table 4-1 presents the default biophysical parameters in LEAF-2 for some important land cover types in the study area. LAI and VFC are maximum leaf area index and vegetation fractional cover; D LAI and D VFC are the maximum decrease in leaf area index and vegetation fractional cover. See

http://www.atmet.com/html/docs/rams/RT1-leaf2-3.pdf for more detailed biophysical characteristics of all land cover types defined in LEAF-2.

Of these biophysical variables, LAI and VFC are assumed to have a simple seasonal dependence (Figure 4-2), which is the function of a cosine distribution, latitude and time in a year. Vegetation is assumed to peak in late July (Julian day = 200) in the northern hemisphere and the reverse in the southern hemisphere. For locations close to the equator, such as a large part of East Africa in this study, LEAF-2 assumes that seasonal variation is reduced to zero. According to Figure 4-2, the built-in spatial and temporal vegetation variations are extremely unrealistic for near-equatorial regions.

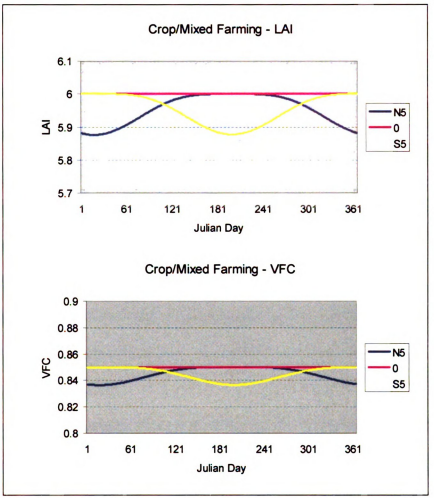


Figure 4-2. Built-in seasonal cycles of LAI (top) and VFC (bottom), illustrated by Crop/Mixed Farming land cover type. "N5" and "S5" refer to 5 degree north and 5 degree south respectively.

In chapter 3, the default land cover dataset in LEAF-2 was found to have significantly lower quality than the newly developed land cover products, such as GLC2000. As previously discussed, there are major shortcomings of the built-in vegetation characteristics in RAMS. This component of study is, therefore, to improve the land surface representation in RAMS by updating the default land cover dataset and the built-in spatial and temporal vegetation dynamics with

GLC2000 and MODIS LAI and VFC products, respectively. Impacts of these improvements will also be investigated.

4.2 Incorporating MODIS LAI & VFC in RAMS

The monthly, 1km MODIS LAI and EVI products for 2003 were downloaded from MODIS group at Boston University ftp://primavera.bu.edu/pub/datasets/MODIS/ (see Chapter 3 for more details). These images were transformed to a Polar Stereographic projection centered at (33°E, 2°S), which corresponds to the projection of the RAMS domain. The LAI data were converted to binary files so that RAMS can read them directly.

It needs to be mentioned that MODIS LAI has a meaning different than that in RAMS. As is standard in the remote sensing community, MODIS LAI is defined as "the area of green leaves per unit area of ground" (Curran 1983, Price 1992), which is sometimes referred to as "effective" LAI. In contrast, LAI in LEAF-2 (Table 4-1) and other land surface models is defined as "the number of leaf layers over the vegetated part of the pixel" (Dickinson, personal communication), which is also referred to as "clump" LAI (Choudhury et al. 1994). Therefore, MODIS LAI was divided by MODIS VFC so that the ingested LAI has the same meaning as defined in RAMS.

VFC data were developed from the 1km monthly MODIS Enhanced Vegetation Index (EVI) product (see Chapter 1 for more description). Compared to the traditional Normalized Difference Vegetation Index (NDVI), EVI takes full advantage of MODIS' state-of-the-art measurement capabilities and has much improved quality (Huete et al. 2002). One noticeable advantage of EVI is that it does not become saturated as easily as NDVI when viewing tropical rainforests and other areas of the Earth with large amounts of green biomass.

Calculating fractional cover from vegetation indices is based on the theory of "Mosaic Pixel", which assumes that a remote sensing pixel has a patchy (mosaic) structure (Kerr et al. 1992, Valor and Caselles 1996). A quantity measured by satellite (ϕ, e.g. vegetation index) for a pixel can be interpreted as the sum of linear contributions from the vegetated area (f_v) and bare soil (1 - f_v):

$$\phi = \phi_V f_V + \phi_S (1 - f_V) \qquad (4\text{-}1)$$

where the subscripts *v* and *s* denote values over fully vegetated and bare soil areas. From this equation, a simple formulation for fractional cover f_v can be derived as:

$$f_V = \frac{\phi - \phi_S}{\phi_V - \phi_S} \qquad (4\text{-}2)$$

While some studies have suggested a non-linear relationship between vegetation fractional cover f_v and vegetation indices ϕ (e.g., Myneni and Williams 1994, Carlson and Ripley 1997), others found the linear relationship is sufficient

(Wittich 1997, Gutman and Ignatov 1998). For this study, the linear relationship is assumed, and Equation (4-2) was used to calculate VFC from MODIS EVI.

EVI for fully vegetated (ϕ_V) and bare soil (ϕ_S) areas needs to be derived in order to use Equation (4-2). Some investigators have suggested that ϕ_V and ϕ_S are dependent on vegetation and soil types, season and geographic regions (e.g., Price 1992, Huete et al. 1994, Zeng, et al. 2000). This study is focused on the East Africa region rather than the whole globe, so that the problem is minimized. In addition, Gutman and Ignatov (1998) found that there are many land cover types for which it may be difficult to find these two constants during the course of a whole year. For example, only one of these two constants can be derived for deserts and evergreen forests. Following the study by Gutman and Ignatov (1998), ϕ_V and ϕ_S in this study are prescribed as seasonally and geographically invariant constants, which correspond to the yearly maximum EVI of the Congo Forest (0.86) and minimum EVI of deserts (0.05) in northern Kenya (Figure 1-1).

Monthly MODIS LAI and VFC were then interpolated linearly to determine daily values so that RAMS can update LAI and VFC values at each time step. It is assumed that monthly MODIS LAI and VFC are observed in the middle of each month.

To ingest the MODIS LAI and VFC data properly into RAMS, the parallel processing architecture must be considered. RAMS is structured in a standard master-node configuration, where the master process handles initialization and output, while the nodes are the main workers. In this configuration, model domain is horizontally decomposed into rectangular subdomains, which cover different sets of grid points and a surrounding boundary region. The nodes work on subdomains and exchange information at the subdomain boundaries. As lower boundary information, LAI and VFC are not assigned during the initialization of RAMS, but are updated dynamically during the simulation. Therefore, the MODIS LAI and VFC data must be assigned to the proper nodes for smooth and correct data incorporation in RAMS.

Figure 4-3 is a schematic illustration of domain decomposition for four-node (I, II, III, and IV) computing in RAMS. The whole domain is 40 (x) by 30 (y) grid cells, which is the domain configuration in this study (This will be further discussed in Section 4.5). Dark areas are boundaries and overlaps between nodes. They exchange information and are coordinated by the master process during each model time step.

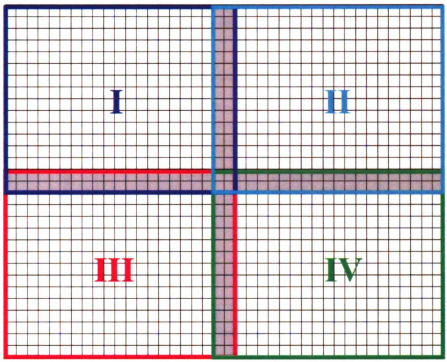


Figure 4-3. Illustration of domain decomposition with four nodes: I, II, III, and IV.

4.3. LAI and VFC comparison

In this section, the built-in and new LAI and VFC are compared both temporally and spatially to illustrate the improved representation of the land surface in the RAMS model.

4.3.1 Temporal comparison

Seasonal LAI and VFC were compared for individual land cover types. Wooded grassland (Figure 4-4) and crop/mixed farming (Figure 4-5) are

illustrated here. The wooded grassland is located at (8.9°S, 23.8°E) (latitude and longitude), while the crop/mixed farming is at (6.2°S, 32.8°E).

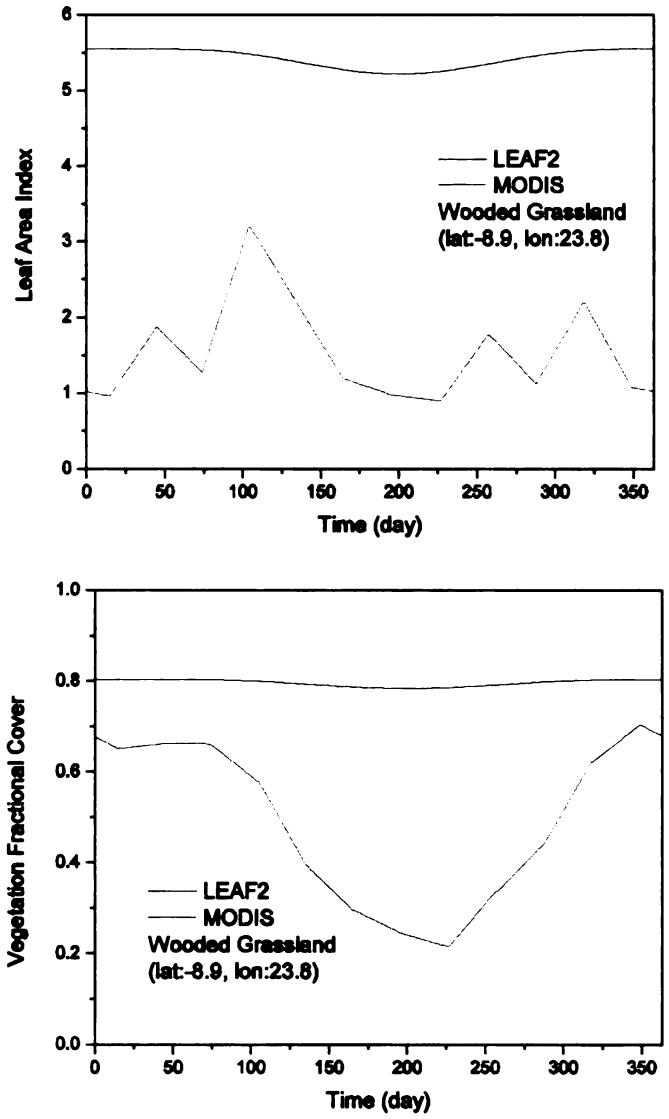


Figure 4-4. Seasonal variations of leaf area index (top) and vegetation fractional cover (bottom) for wooded grassland. The black lines are LEAF2 built-in variations, and the red lines are from MODIS products.

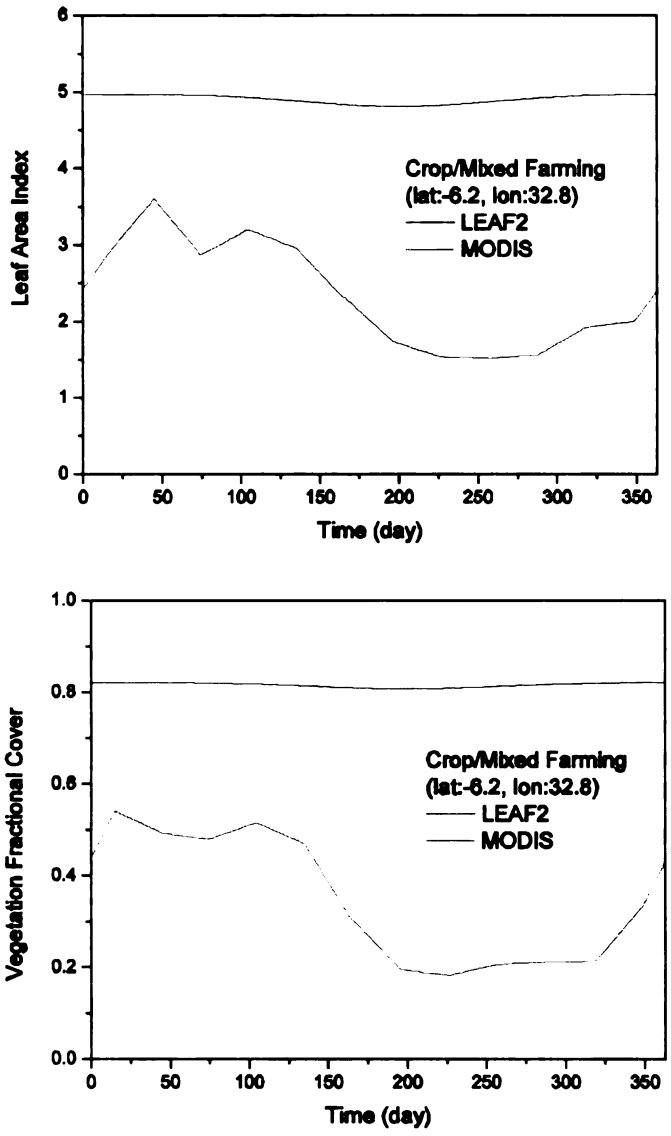


Figure 4-5. Same as in Figure 4-4 but for crop/mixed farming.

According to these two figures, RAMS built-in LAI and VFC displayed little annual variation and had much higher values compared to the MODIS observations. As discussed previously, the seasonal variation of vegetation in RAMS is assumed to have a cosine distribution, which varies with latitude (see Figure 4-2). The focus area in this study straddles the equator and has low latitude. As a result, the built-in seasonal vegetation cycles in RAMS have unrealistically small temporal variation.

4.3.2 Spatial comparison

Figure 4-6 and Figure 4-7 present the spatial comparison of LAI and VFC respectively for different time periods in 2003: March, June and September. In addition to the built-in and improved LAI and VFC, the original 1 km MODIS imagery is also included in the comparison. The resolution for the built-in and improved LAI and VFC maps is 50 km, which corresponds to the RAMS grid spacing (see section 4.5 for model configuration). Colors in the 50 km maps indicate LAI and VFC values for the biggest patches in the grid cells.

Overall, the built-in LAI and VFC are extremely homogeneous spatially. Except for deserts and lakes, vegetation has little variation over the domain. The built-in data show the Congo forest to have similar biophysical characteristics as the semi-arid areas in the east. As shown in the previous section, the built-in LAI and VFC also present unrealistic temporal variation. For example, the observed VFC in a large part of southeast portion of the study area decreases significantly

in the second half of the year, which is likely related to the pattern of ITCZ. However, this is completely missed in the default VFC.

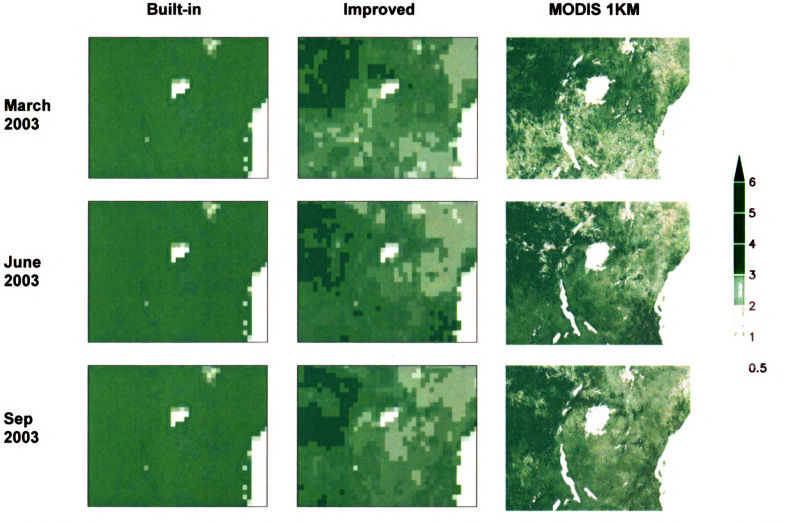


Figure 4-6. Spatial comparison of the built-in, improved and original 1 km LAI for March, June and September, 2003

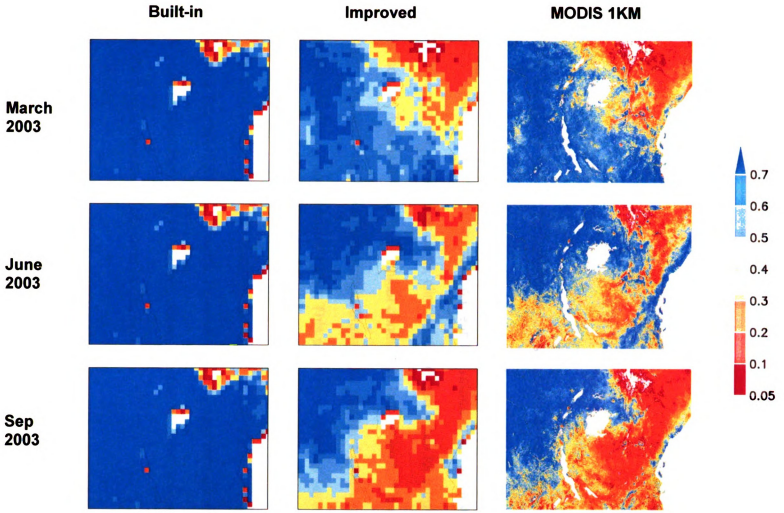


Figure 4-7. Same as in Figure 4-6 but for VFC.

In the following sections, RAMS will be run with the updated land cover dataset (GLC2000) and improved LAI and VFC for the whole year of 2003. The impacts will then be examined on the basis of simulated precipitation and land surface temperature.

4.4 RAMS and configuration

RAMS 4.4 is an atmospheric model which is capable of both numerical weather prediction and regional climate simulation. In a philosophical sense, numerical weather prediction depends on the initial values of the state variables of the atmosphere. On the other hand, climate simulation is run for longer periods of time, so that it is insensitive to the initial conditions, but dependent on boundary conditions such as ocean temperature, land use, and greenhouse gas concentrations (Giorgi and Mearns 1999). This simulation includes some parts of the climate system such as a full treatment of atmospheric dynamics, thermodynamics and moisture processes, along with a Soil-Vegetation-Atmosphere Transfer (SVAT) scheme. However, unlike some climate models, RAMS does not include a fully interactive ocean. Rather, it treats ocean surface temperature as a prescribed boundary condition.

The soil model in LEAF-2 consists of 11 vertical layers spanning a depth of 2.1m, and the soil temperature profile in the initial condition was determined by a deviation from the initial air temperature in the lowest atmospheric level. The soil

moisture content for the top layer was initialized as 35% of the saturation value, which was horizontally homogeneous over the domain. This percentage was increased with depth to a maximum of 55% at 48 cm and below. Moisture flux between soil layers was parameterized in LEAF-2 based on the multilayer soil model described by Tremback and Kessler (1985). Both energy and moisture fluxes between LEAF-2 components (i.e., vegetation, canopy air, and each soil and snow cover layer) are illustrated in detail in Walko, et al. (2000, also refer to section 4.1).

Soil moisture can play an important role in surface-atmosphere interactions particularly through moisture "memory" in semi-arid regions like in Kenya and Tanzania (Fig 3-1). The presence of soil moisture influences the partitioning of latent and sensible heat, thereby affecting the development of shallow convection. However, the soil types in East Africa are poorly mapped, and available soil moisture values for the region are speculative due to data scarcity. It needs to be emphasized that the role of surface parameters, including soil moisture, can strongly affect the model solution (Ducharne and Laval 2000). In the absence of reliable data, and to avoid introducing more complex uncertainties into this experiment, this homogeneous approach was chosen.

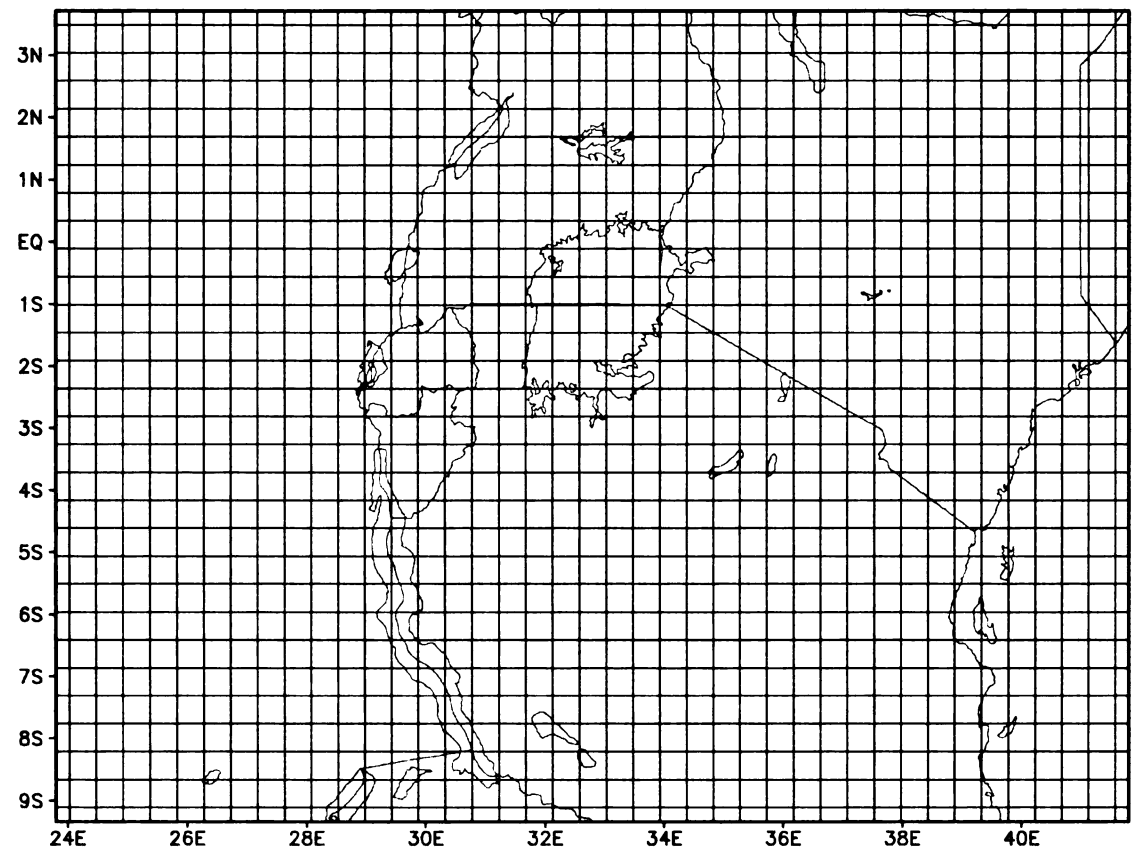


Figure 4-8. RAMS domain with Δx = 50 km

In Figure 4-8, a single model grid with 50 km horizontal grid spacing was used to cover the study area (see Fig 3-1). The horizontal grid spacing was set at 50 km in consideration of the domain size and the computational requirements. For the land surface, the standard RAMS 30-arc sec topography dataset was used. The grid extended over 32 vertical levels, with a layer thickness of 80m near the surface and stretching to 1900m at the top of the domain. The model was driven by 6-hourly lateral boundary conditions derived from the National Centers for Environmental Prediction (NCEP) atmospheric reanalysis product (Kalnay et al. 1996). The model time step was 90 seconds with the output period

set to every six hours. At each time step, the reanalysis data were nudged over five outer grid points.

The radiative transfer scheme of Chen and Cotton (1983) was used to parameterize the vertical flux of shortwave and longwave radiation. Horizontal diffusion coefficients were computed based on the modified Smagorinsky formulation (Smagorinsky 1963), and the vertical diffusion was parameterized according to the scheme of Mellor and Yamada (1982). The bulk microphysics parameterization was activated, which allows the model to consider the effect of moisture in all phases. The sea surface temperature was specified using the 1° monthly climatological dataset from NCEP (Reynolds and Smith 1994).

Three experiments were carried out to evaluate the impacts of the improved land surface parameterization in RAMS. In the first run, here called the "default" (DEF) run, the default land cover dataset (OGE) and the built-in LAI and VFC were used. In the second run, here called the GLC run, the built-in LAI and VFC were used but with OGE land cover replaced by GLC2000. Both GLC2000 and MODIS derived LAI and VFC were used for the third run, which is called GLC+LAI+VFC in this study.

4.5 TRMM precipitation and MODIS land surface temperature

4.5.1 Weather station data

Observed climatic variables, e.g., rainfall and temperature, are usually used for assessing model performance and outputs. The main problem with traditional meteorological observations is their poor representation of the grid-point average simulated by a model. Well documented station observations of climate variables are mainly located in populated and industrialized regions. Observations are very sparse both spatially and temporally over many regions of the world, which brings inevitable challenges for climate modeling studies.

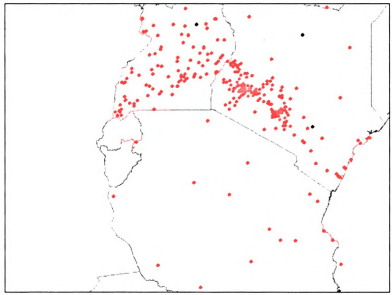


Figure 4-9. Distribution of weather stations in the study area

Figure 4-9 shows the weather stations in the study area. Spatially, there are very few stations in Congo, Tanzania, and east Kenya. More importantly, many of the stations shown here do not have a continuous time series of

observations. To solve this problem, climatic variables observed by satellites are used in this study: rainfall data from the TRMM satellite and land surface temperature from MODIS.

4.5.2 TRMM precipitation

Tropical Rainfall Measuring Mission (TRMM) is a joint US – Japanese satellite that was launched in November 1997 (Simpson et al. 1998, Kummerow et al. 2000). The primary mission of TRMM is to measure precipitation in the Tropics where conventional surface observations are scarce. TRMM provides the most accurate global tropical rain estimates to date by using a unique combination of instruments (the TRMM Microwave Imager and the Precipitation Radar) designed specially for rain observation and by using a low-inclination (35^{o}) orbit to provide excellent coverage of the Tropics (Kummerow et al. 2000). The TRMM products have recently been evaluated over East Africa using station observations (Dinku et al. 2007), and they perform reasonably well. Although the primary objective of TRMM was to improve climate models and to aid them in climate prediction (Kummerow et al. 2000), very few studies have fully utilized this dataset for model validation and rainfall data assimilation.

For this study, the latest TRMM product (3B42 version 6) is used. Its primary characteristics are listed in Table 4-2 (http://daac.gsfc.nasa.gov/precipitation/TRMM_README).

Table 4-2. TRMM 3B-42 characteristics

Temporal Coverage	Start Date: Jan 1, 1998; Stop Date: -
Spatial Coverage	Latitude: 50°S – 50°N; Longitude: 180°W – 180°E
Temporal Resolution	3-Hourly
Spatial Resolution	0.25° × 0.25°
Average File Size	Original: ~ 4.5MB
File Type	HDF

3B42 V6 data with global coverage for 2003 were downloaded from http://lake.nascom.nasa.gov/Giovanni/tovas . These data were subset and compared with the 3-hour precipitation file produced by RAMS. An online interface for visualization and analysis of TRMM data is also available from this website.

4.5.3 MODIS land surface temperature

Land surface temperature (LST, also referred to as radiometric temperature or skin temperature) is inferred from the thermal emission of the land surface and is more directly related to the surface properties than the surface air temperature. For bare soil, LST is the soil surface temperature. For vegetated areas, LST is generally some average of the temperature of various canopy and soil surfaces (Qin and Karnieli 1999, Jin et al. 1997). It is one of the

key parameters in the physics of land surface processes on regional to global scales. For example, sensible heat flux is proportional to the difference between instantaneous LST and the temperature of the overlying air. Many researchers have used LST to estimate the sensible heat flux as well as soil moisture (Sun and Mahrt 1995, Wang 1992).

Before satellites, LST could only be obtained by ground-based or airborne radiometers, which have very limited capabilities to provide convincing regional to global views. In the late 1970s, global measurements of LST derived from AVHRR became available. The split-window algorithm (Price 1984, Becker 1987), which was originally devised for sea surface temperature (SST), is the most extensively used methodology to drive LST from the two thermal bands of AVHRR: band 4 (10.5 – 11.5 um) and band 5 (11.5 – 12.5 um) (Qin and Karnieli 1999). A major problem in using split-window methods is that the surface emissivities are required with accuracy better than 0.01, which are almost impossible to estimate for many land cover types such as semi-arid and arid areas (Wan and Li, 1997).

With multiple thermal-infrared bands designed for LST retrievals, MODIS is providing a much more accurate LST product for both regional and global climate modeling. A physical-based day/night algorithm (Wan and Li 1997) was developed to retrieve LST at 5 km resolution from a pair of daytime and nighttime MODIS data in seven thermal bands (bands 20, 22, 23, 29, and 31 – 33).

Compared with *in situ* measurements, the MODIS LST accuracy is better than 1K in the range from 263K to 300K (Wan et al. 2004).

More importantly, MODIS is capable of providing diurnal LST observations. Two MODIS instruments (Salomonson et al. 1989) have been launched. The first was launched on the Earth Observing System (EOS) morning platform Terra (called MOD in this study), and the second was launched on the EOS afternoon platform Aqua (called MYD in this study). The Terra overpass times are around 10:30 (local solar time) and 22:30, while the Aqua overpass time is around 13:30 and 1:30. For this study area, 10:30, 22:30, 13:30, and 1:30 in local solar time correspond to approximately 08:00, 20:00, 11:00, and 23:00 in Coordinated Universal Time (UTC). MOD products have been made available to the public for about five years, while MYD became available only recently. Despite four observations per day, MOD and MYD combined are now providing unprecedented diurnal LST information. In addition, MYD measurements are obtained at 11:00 UTC which are more close to the time of maximum temperature of the land surface. It is therefore more suitable for regional and global change studies.

AVHRR LST has been compared with that simulated in a global climate model (Jin et al. 1997). In this study, the newly developed MOD and MYD are used for the first time to evaluate the LST calculated in a regional climate model

(RAMS) and to study the impact of changing the land surface representation. Diurnal LST produced by RAMS is also examined.

RAMS does not output LST directly. In this study, canopy temperature and top-layer soil temperature are combined to calculate LST in each grid cell on the basis of vegetation fractional cover (VFC), using the following equation:

$$LST^4 = \sigma_V T_V^4 + (1-\sigma_V) T_g^4 \qquad (4\text{-}3)$$

where σ_V is VFC, T_V is the canopy temperature, and T_g is the top-layer soil temperature (Jin et al. 1997). From equation (4-3), vegetation density (leaf area index) and vegetation fraction can have a significant impact on simulated LST. By incorporating the more spatially and temporally realistic MODIS LAI and VFC, RAMS is expected to produce LST with improved characteristics.

4.5.4 LST data quality

Monthly MOD and MYD LST data (version 4) with 0.05° spatial resolution was downloaded from the EOS Data Gateway (http://edcimswww.cr.usgs.gov). They were subset and resampled to 50 km resolution to match with RAMS output. It needs to be pointed out that MODIS LST is valid only under clear-sky conditions. For this low-latitude study area (Fig 3-1), the daily product is usually severely contaminated by clouds. Monthly MODIS LST is thus used in this study. It is composited and averaged on the basis of clear-sky observations. However, monthly LSTs in several months of 2003 were still adversely affected by the

persistent cloud cover, especially over the western part of the study area (Congo forest) during nighttime. Figure 4-10 is an illustration of cloud contamination in the original (0.05° spatial resolution) nighttime MYD LST for November, 2003. Grey pixels have valid LST values, while yellow pixels indicate invalid LST observations, which include the ocean and a portion of Congo forest.

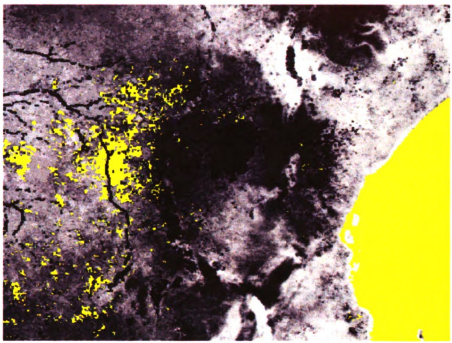


Figure 4-10. Cloud contamination in nighttime MYD LST for November, 2003.

The data quality of MOD and MYD LST in 2003 is summarized in Table 4-3, where √ indicates no significant cloud cover and X indicates significant cloud contamination. It is noticeable that most cloud contamination occurs during night time and in the raining seasons.

Table 4-3. Data quality of MODIS LST in 2003

	MYD		MOD	
Month	Day	Night	Day	Night
1	√	X	√	√
2	√	√	√	√
3	√	√	√	X
4	√	√	√	√
5	√	√	√	√
6	√	√	√	√
7	√	√	√	√
8	√	√	√	√
9	√	X	√	√
10	√	X	√	X
11	√	X	√	X
12	√	X	√	X

In this study, daytime MYD (11:00 UTC) is assumed as the maximum daily LST and is compared with RAMS LST at (12:00 UTC) for the whole domain. Nighttime MYD (23:00 UTC) is assumed as the minimum daily LST, recognizing that minimum temperature may actually occur a couple of hours later in the early morning. From 0:00 UTC, RAMS outputs after every three hours. Thus, there is a one hour difference between MOD and MYD observing times and the simulation times in RAMS. These time differences are assumed to have a negligible influence on the comparison in this study. Since nighttime LST is usually contaminated with clouds in the western part of the domain (Fig 4-10), the minimum LST comparison is restricted to the eastern part of the domain. In

addition, diurnal LST is compared from May to August, due to minimum cloud contamination in this time period.

4.6. Results

In this section, LST and precipitation in 2003 simulated by the three RAMS runs (DEF, GLC, and GLC+LAI+VFC) are compared spatially and temporally to examine the effects of the improved land surface representation. They are also evaluated against MODIS and TRMM products.

4.6.1 LST

4.6.1.1 Temporal comparison

Figure 4-11 shows the RAMS simulated maximum daily LSTs and the MYD monthly composited observations, which are averaged for the whole study area excluding water bodies. The three runs are differentiated in colors and bold lines are simply 30-day averages of daily results. The impacts of the improved land surface conditions are significant, especially when MODIS LAI and VFC products are used. In contrast with the RAMS default land surface, LST seasonal variation has been dramatically improved in GLC+LAI+VFC (green line in Figure 4-11). According to MYD observations the lowest LST in 2003 occurs in May. This is correctly simulated by GLC+LAI+VFC, but is incorrectly simulated to be in July by DEF and GLC runs. LSTs peak at about the same time periods (March

and September) in all three runs. However, RAMS strongly underestimates LST in the second half of the year, probably as a result of its unrealistic LAI and VFC information. Overall, GLC2000 increases the maximum daily LST by about one degree, but fails to change the seasonal dynamics. LST flattens out later in the year, both with and without GLC2000.

As discussed previously, the land surface is primarily represented by both land cover types and their related biophysical parameters (see Table 4-1). Introducing a new land cover dataset, GLC2000, provides a better description of the spatial distribution of land cover types across the study domain. This should help to produce better spatial characteristics of LST (discussed later). However, GLC2000 is not able to modify the seasonal dynamics which are still prescribed as simple mathematical equations in LEAF-2 (Figure 4-2). In Equation (4-3), the temperatures of vegetation and bare soil as well as their relative contributions in a grid cell, vary not just by location, but also over time. As a result, using GLC2000 alone does not improve the LST seasonality.

By incorporating temporally explicit MODIS LAI and VFC (Figs. 4-6 and 4-7) RAMS improves its performance of simulating the seasonal dynamics of LST in two ways. Informed by the MODIS LAI data, a more realistic vegetation density of the vegetated area in a grid cell may help to improve the interactions of plants and overlying atmosphere. The other one, which is more important, is the relative contributions of vegetation and bare soil as determined by VFC. Significantly

increased areal extent of bare soil in the eastern domain during the second half of the year (Figure 4-7) may help to lift the flat LST curves after June (Figure 4-11).

It is noticeable that the MODIS-observed LST is much higher than the RAMS simulation for a large part of the year. Remember that monthly MODIS LST is composited on the basis of clear-sky observations. By screening out cloud contamination, the MODIS observations are biased toward the highest values in a month, since clouds diminish solar insulation and thus cause much lower LSTs.

Minimum daily LST is also compared in Figure 4-12. Because a significant amount of invalid LST pixels exist in the nighttime MYD in the western domain from September to January (see table 4-3), this comparison was conducted only for the eastern domain, which is approximately the eastern half of the study area. Water bodies were excluded in this comparison as well. Generally, RAMS produces much better minimum LST dynamics in all three runs. This is possibly due to less moisture in this arid/semi-arid area. Unlike the maximum daily LST, the lowest minimum LST occurs in July. Seasonal variation is as large as about four degrees. Similarly as in Figure 4-11, RAMS produces lower LSTs at the beginning of the year, but higher LSTs after June using the GLC+LAI+VFC run. In addition, GLC2000 does not have much impact on the LST seasonality. Also, MODIS observations are lower than RAMS which is similarly due to MODIS'

clear-sky observations. The surface is usually cooler at night with less cloud cover.

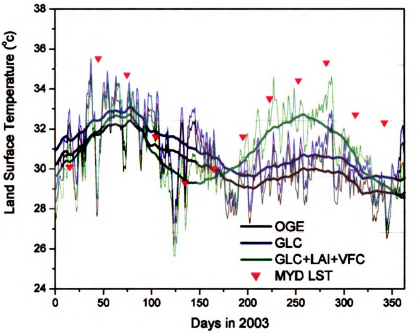


Figure 4-11. Temporal comparison of maximum daily LST from three RAMS runs: OGE, GLC, and GLC+LAI+VFC, and monthly MYD observations for the whole study domain.

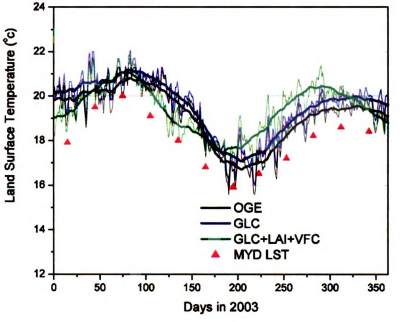


Figure 4-12. Same as in Figure 4-11 but for minimum daily LST and for eastern domain only

4.6.1.2 Spatial comparison

In Figure 4-13, the maximum daily LSTs from MODIS (MYD daytime) and RAMS are spatially compared month by month. Images on the top are from the MODIS observations, resampled to 50 km resolution. Only LSTs from the DEF run (middle) and the GLC+LAI+VFC run (bottom) are included. Because the comparison is in terms of land surface temperature all water bodies (lakes and oceans) were masked out (white color). The magnitude of LST is indicated by color: yellow to red show high LSTs while green to blue show relatively low LSTs.

MODIS LSTs peak from February to March and again from August to September, and reach their minimum around May. This is consistent with the bimodal temporal pattern shown in Figure 4-11. This feature is obviously captured by the GLC+LAI+VFC run, but is completely missed by the DEF run. LSTs do not vary much after June for the DEF run. Spatially, the western domain which is covered by more vegetation and has more moisture and rainfall (discussed later) has lower LSTs compared to the eastern part. Both the DEF and the GLC+LAI+VFC runs seem to capture this feature, but the latter captures this contrast much better. More importantly, MODIS-observed LSTs in the east shows a strong ITCZ related pattern, in which high LSTs migrate from north to south with time. This feature is fully captured by the GLC+LAI+VFC run. In Figure 4-7, the MODIS VFC shows a similar pattern. This confirms the importance of VFC in calculating LST as is manifested in Equation (4-3).

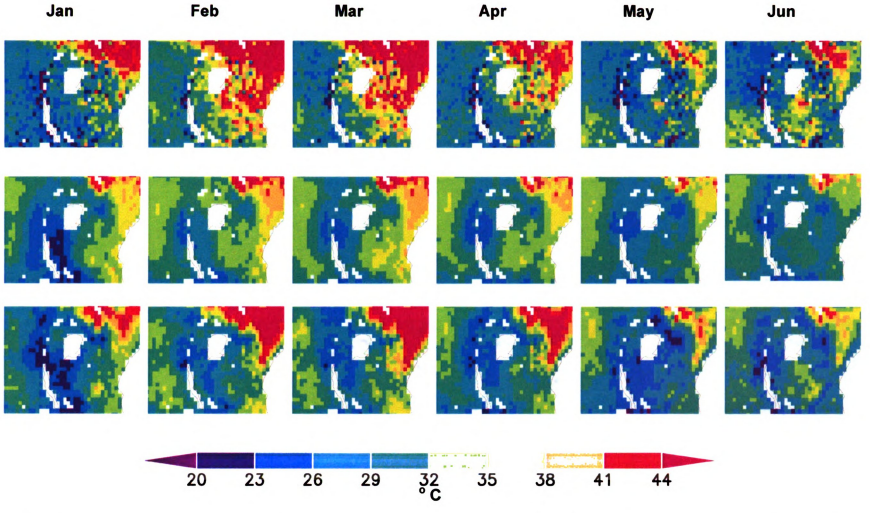


Figure 4-13. Spatial comparison of LST from MODIS observations (Top), RAMS with default land cover and biophysical parameters, DEF run (Middle), and RAMS with GLC2000 and MODIS LAI/VFC, GLC+LAI+VFC run (Bottom)

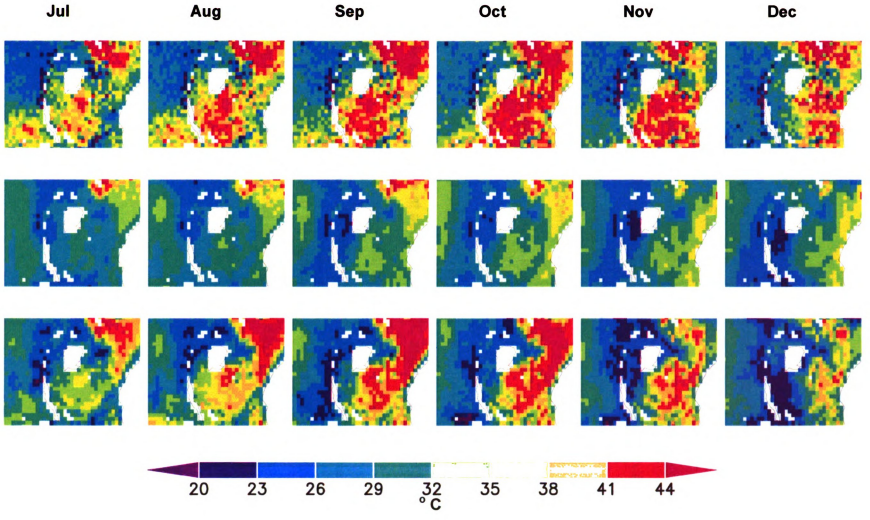
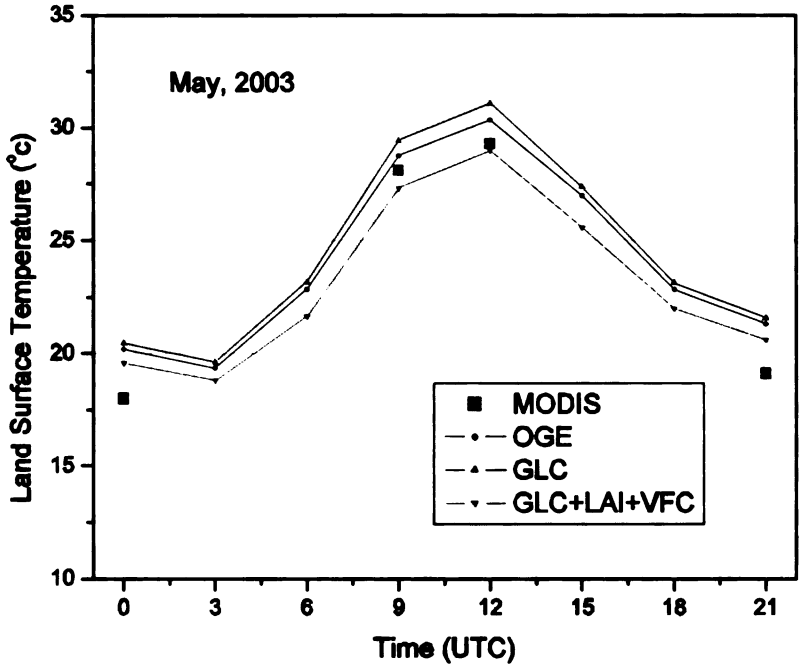


Figure 4-13. *continued*

4.6.1.3 Diurnal LST

The impacts of the land surface representation in RAMS are further examined by looking at diurnal LSTs, taking advantage of the multiple daily observations by MODIS Terra (MOD) and MODIS Aqua (MYD). Due to cloud contamination (see table 4-3), only four months, May to August, are analyzed. In Figure 4-14, the RAMS simulated LSTs, which are output eight times per simulated day (every three hours), are shown by the color lines, while the four MODIS observations are shown in red squares. Time is in UTC, which is about three hours different from local time in the study area.

This figure shows that RAMS captures the diurnal cycles quite well. LST reaches its daily maximum at about 12:00 UTC and the daily minimum at about 03:00 UTC. As shown in Figure 4-12, RAMS produces higher LST at night (0:00 and 21:00 UTC). During the day (09:00 and 12:00 UTC), the simulated LST is close to the MODIS values in May and June. However, in July and August the differences between simulated and observed daytime LSTs magnify. This is also shown in Figure 4-11. During the day in July and August, only the LSTs simulated by the GLC+LAI+VFC run are close to the MODIS observations.



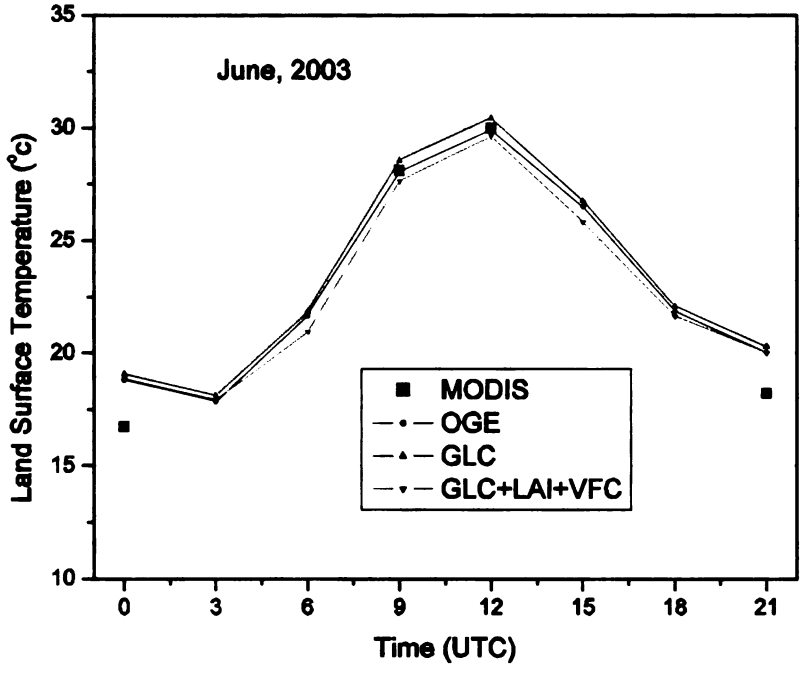
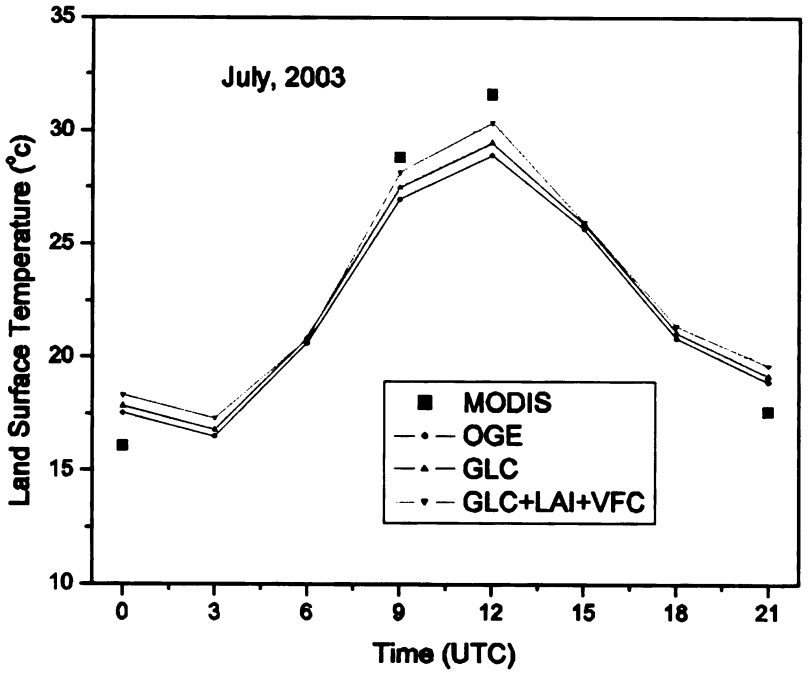


Figure 4-14. Monthly averaged diurnal cycles from MODIS and three RAMS runs: OGE, GLC, and GLC+LAI+VFC



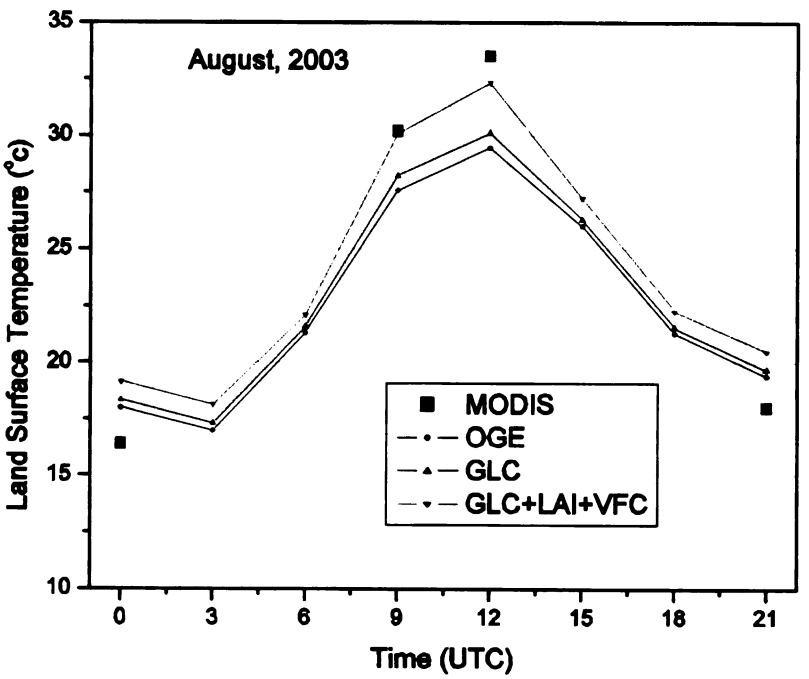


Figure 4-14. *continued*

4.6.2 Precipitation

Figures 4-15 and 4-16 show the TRMM-observed 3-hour precipitation rates and accumulated precipitation in 2003, respectively. These TRMM observations are compared with the three RAMS simulations: DEF, GLC, and GLC+LAI+VFC. RAMS is able to produce the general temporal and spatial dynamics of rainfall in this area considering RAMS has a much lower spatial resolution than the TRMM observations (50km vs. about 27km). The two wet seasons separated by a dry season in-between (around June) are clearly captured in Figure 4-15. Spatially, the major features are captured by RAMS, such as high rainfall areas over the Congo forest and Lake Victoria and relatively dry areas in eastern Kenya and Tanzania. However, RAMS produces little rainfall at the left and right boundaries, which maybe due to the boundary nudging employed in the model. However, as shown in Figure 4-15, RAMS underestimates the major wet season (March – May), but overestimates the second wet season (about October to November) in 2003. In addition, RAMS tends to generate a larger contiguous area of high rainfall over the Congo forest compared to the TRMM-observed, three insular areas of high rainfall in this area.

In contrast to LST, the RAMS simulated precipitation seems to be less sensitive to different land surface representations. All three experiments produced similar spatial and temporal rainfall patterns. In Figure 4-16, however, the GLC+LAI+VFC run produced a somewhat more realistic compact high rainfall area over the Congo forest compared to the other two experiments.

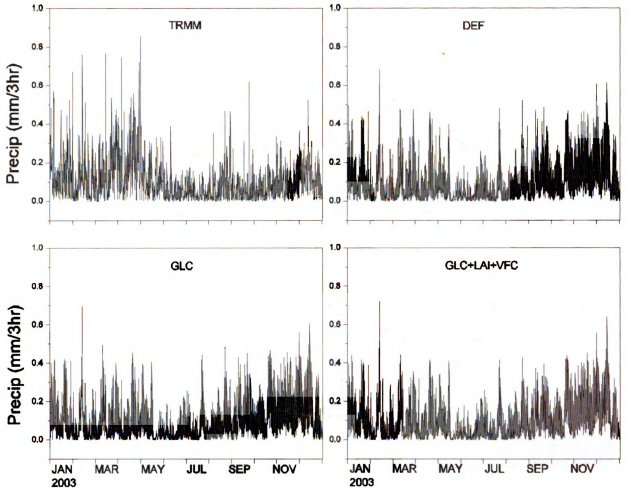


Figure 4-15. Domain averaged three-hour precipitation rate from TRMM for 2003 and three RAMS runs: DEF, GLC, and GLC+LAI+VFC

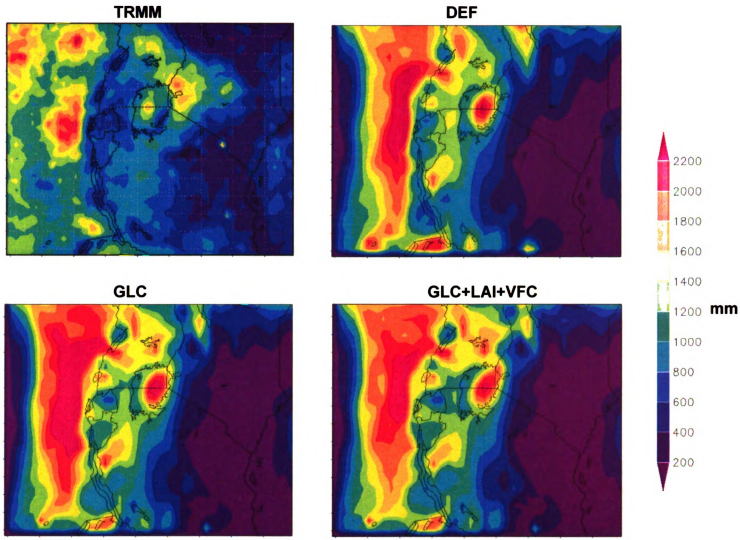


Figure 4-16. Accumulated precipitation in 2003 from TRMM and three RAMS runs: DEF, GLC, and GLC+LAI+VFC

4.7 Conclusions and discussions

In RAMS, the default land surface is represented by an outdated land cover dataset and a set of prescribed parameters. Vegetation phenology is unrealistically simplified by mathematical equations, which primarily are functions of temperature and latitude. In comparison to the corresponding MODIS LAI and VFC products, the spatial and seasonal vegetation dynamics of East Africa are poorly represented by the default approach in RAMS. Overall, the default LAI and VFC in RAMS are too homogeneous among different vegetation types across the domain and are almost invariant over the year, which are very uncharacteristic of the real world situation.

In this study, the GLC2000 land cover product, which was evaluated in Chapter 3 to have the best quality, and newly developed spatially and temporally explicit MODIS LAI and VFC products are directly ingested into RAMS to improve the representation of the land surface. Their impacts on regional climate simulations were examined by three different RAMS experiments: DEF (default OGE land cover and the built-in LAI and VFC), GLC (OGE is updated by GLC2000 but with the built-in LAI and VFC), and GLC+LAI+VFC (GLC2000 plus the MODIS LAI & VFC are used). Results show that the temporal and spatial LST dynamics from the GLC+LAI+VFC run are greatly improved over the other two experiments. Temporally, incorporating MODIS LAI and VFC enables RAMS to capture the bimodal LST characteristics, especially the peak from August to

October in 2003. Spatially, RAMS' capability to simulate ITCZ-related LST migrations in the eastern domain is greatly improved. GLC2000 alone, however, was not sufficient to improve the LST seasonal variation, but did change the overall magnitude by about one degree.

In temporal terms, RAMS produces precipitation fairly well compared with precipitation observations from the TRMM satellite. Two wet seasons and the relatively dry period in 2003 were reasonably captured. Spatially, the general major high rain features in the western domain and the dry areas in the eastern domain are very well produced. Unlike LST, precipitation in RAMS seems to be less sensitive to different land surface conditions. Rainfall outputs from the three RAMS experiments: DEF, GLC, and GLC+LAI+VFC have closely similar temporal and spatial characteristics as examined by the precipitation rate over time and the accumulated rainfall in 2003.

The improvement of the daytime LST can be attributed to more realistic description of the surface evapotranspiration process. The MODIS biophysical products, especially VFC, capture the ITCZ related less-vegetated zone in the eastern domain during the later part of the year (Figure 4-7). There is thus less moisture evapotranspiration activity. The surface temperature in GLC+LAI+VFC therefore increases compared to DEF and GLC with unrealistically covered vegetation. Precipitation in RAMS is governed by large-scale fields and model

parameterizations. The improvement of the land surface representation is not sufficient to change the simulated precipitation significantly.

It is worth mentioning that climatic variables from satellite observations can help to improve the climate modeling studies in areas with sparse spatial and temporal station measurements, as demonstrated in this study. TRMM precipitation retrievals with three-hour temporal resolution is an excellent validation variable for regional climate modeling. Though it was designed for climate modeling studies, very few published works use these data except the regular comparison with traditional data conducted outside of the climate modeling community. Furthermore, the land surface temperature measured by satellites has been shown to be very useful in GCM studies. For the first time, this study compared the model-simulated LSTs with the newly developed MODIS Terra and MODIS Aqua LSTs in a regional model (RAMS). Spatially and temporally explicit, as well as diurnal, LST information is fully utilized in the RAMS model in this study.

Chapter 5

Impacts of Land Cover Classification Accuracy on Regional Climate Simulations

5.1 Introduction

Human activities are transforming the surface of the Earth at an accelerated pace. Such disturbance of the land can affect the local, regional, and global climate by changing the energy balance on the Earth's surface and the chemical composition of the atmosphere (Chase et al. 1998, Houghton et al. 1999, Pielke 2001). Over the past few decades, land use/cover has been widely recognized as a critical factor mediating socioeconomic, political and cultural behavior and global climate change (IGBP 1990, Lambin et al. 1999, Watson et al. 2000). Numerous attempts have been made to understand past climate changes and to project potential future climate changes by incorporating reconstructed historical land cover changes and projected possible future land cover changes into numerical simulations (Xue 1997, Pielke et al. 1999, Chase et al. 2000, DeFries et al. 2002, Taylor et al. 2002). Recent studies have suggested that land use/cover change has a first-order climate effect at the global scale (Feddema et al. 2005).

No land cover dataset is one hundred percent accurate, even if developed from the most advanced satellite images. Other factors, such as the classification

method, the sample size of evaluation data, and the inherent subjective characteristics of classification, can increase the uncertainties contained in land cover datasets. Such limitations have been recognized in the remote sensing community, and therefore quantitative accuracy assessment has been emphasized in most recent land cover classification research (Foody 2002). Some target accuracy thresholds have recently been recommended in an attempt to provide guidelines to the classification quality. Thomlinson et al. (1999), for example, set as a target an overall accuracy of 85% with no class less than 70% accurate. However, classification accuracy is usually interpreted differently from the viewpoint of various users. The effect of land cover accuracy for a particular application, such as climate modeling in this study, remains an unanswered question. The accuracy targets commonly specified have largely not been tested from the perspective of the operational use of land cover data.

The objective of this chapter is to examine how the classification accuracy of a land cover dataset employed in a land surface scheme affects simulated cumulative precipitation in a regional climate model. (Here, "regional climate model" means a limited area model with high resolution, generally with grid spacing less than 100 km, run for a simulation time of more than approximately two weeks' length, so that the initial atmospheric conditions have been forgotten (Jacob and Podzun 1997).) The hypothesis of this study is that degradation of land cover classification accuracy will not result in a significant change in simulated regional climate until it reaches a certain threshold. By identifying this

threshold, the requirement of classification accuracy and the uncertainty originating from the land cover classification in the regional climate simulation analysis can be determined.

In addition, three follow-on experiments were conducted to investigate how certain model parameterizations influence this effect. The parameterizations examined in this study are the convection schemes and interior nudging, which have been shown to influence the atmospheric response to surface boundary forcing (Weaver et al. 2002, Castro et al. 2005). These follow-on experiments help to illustrate how land classification error can propagate to factors that govern precipitation in the climate model.

5.2 Methodology

The RAMS model is utilized to simulate the main wet season in East Africa from March to May for the year 2003. Unlike in Chapter 4, only three months are simulated in this chapter because the ensemble run in this study requires significantly more computing resources. The study area is shown in the Figure 3-1. To better represent the land surface characteristics, the default land cover in the model (OGE) is replaced by the improved GLC2000, which ranked the highest in the biophysical evaluation in Chapter 3. Based on GLC2000, classification error with increasing magnitude is then simulated. Cumulative

precipitation from simulations with different classification accuracies is then examined.

The RAMS model and its configuration in this study have been described in detail in Section 4.4 of Chapter 4. In the basic experiment, the Kain-Fritsch (KF) convection scheme (Kain and Fritsch 1993) was used with no interior nudging. In the three follow-on experiments, the effects of a different convection scheme by Kuo (1974) and interior nudging are explored.

5.3 Land cover dataset

For an updated representation of the land surface cover, the GLC2000 dataset was used in these experiments, replacing the default OGE file. In order to be able to use the biophysical parameters adopted from the Biosphere-Atmosphere Transfer Scheme (BATS) (Dickinson et al. 1986), the GLC2000 classes were cross-referenced based on the results of multiple assessments (Torbick et al. 2006). The predominant five non-water land cover types after cross-referencing are presented in Table 4-1, with the most important biophysical parameters listed. Combined, the five predominant types comprise 62.1% of the total area, while ocean and inland water combined comprise 12.3%. The largest inland water body in this area is Lake Victoria in the center of the model domain (Figure 3-1). In the default LEAF-2 methodology, the original 1 km land cover data were sampled to reduce the demand on computing resources used to

initialize the model. Only one pixel's value, for example, is taken from a 5×5 pixel block for a configuration of a 50 km horizontal spacing. As a result, details of the input land cover are lost. In this study, detailed land cover input is needed, and therefore the sampling strategy was modified to take every 1 km land cover pixel in a grid cell.

5.4 Land cover accuracy

Land cover accuracy is commonly defined as the degree to which the derived classification agrees with reality (Foody 2002). Here, classification error at the1 km level was simulated as a random difference from GLC2000 (Figure 5-1a), the initial baseline land cover which was assumed to be 100% accurate. Specifically, random locations in the 1 km GLC2000 were selected, and the original land cover type at each of these selected locations was replaced by a type randomly chosen from the five predominant types (Table 4-1). Only the land cover types could be randomly altered since in practice it is less likely that water bodies are misclassified. The five predominant land classes were chosen, because it is reasonable to assume that they have more chance to be misclassified than less abundant classes.

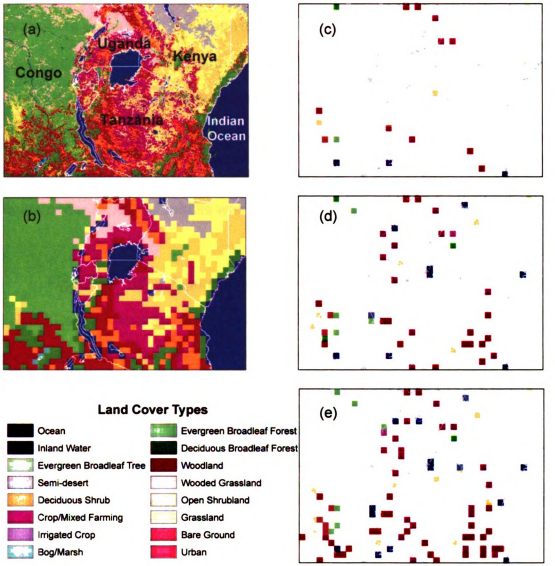


Figure 5-1. Cross referenced GLC2000 with 1km resolution (a) and 50km resolution (b), and simulated land cover classification errors: 10% (c), 30% (d), and 50% (e). Land cover types in (b), (c), (d), and (e) only represent the biggest patches in grid cells. See texts for more details.

Classification errors with magnitudes ranging from 5% to 50% at 5% intervals were generated. The magnitude of error is determined by the proportion of converted pixels in the 1 km GLC2000. Fifty percent error was the maximum level tested as it is assumed that most land cover products could reach 50% accuracy levels. These 1 km land covers with degraded classification accuracies were used to initialize the land surface in sequential RAMS simulations, and the behavior of the simulated results was examined.

In Figure 5-1, 10% (Figure 5-1c), 30% (Figure 5-1d), and 50% (Figure 5-1e) classification errors are presented. For the sake of clarity, only the most predominant patches in each 50 km RAMS grid cell are illustrated because simulated errors and their gradual increase would be hard to see at a 1 km resolution (Figure 5-1a). Figure 5-1b presents the land cover at a 50 km resolution, assumed to be 100% accurate, with each grid cell showing only its most predominant cover type. Figures 5-1c, d, and e show those model grid cells with the predominant land cover changed following the introduction of random classification errors.

Despite random selection at the 1 km resolution, the errors do not appear to be distributed randomly over the domain when viewed at 50 km level. Instead, they tend to occur at the transition zones between major types (Figure 5-1b), where it is likely that two land cover types are approximately equal in frequency within the grid cell. Converting a few pixels may alter which land cover type is the

predominant patch. For example, most changes in evergreen broadleaf forests in Figures 5-1c, d, and e occur at the edge of the Congo forest. For grid cells with strongly dominant types, such as the Congo forest, random errors are less likely to change the dominance of the biggest patch. Transitions to woodland appear to have a higher frequency than do the other four types (see especially in Figure 5-1e). This is due to the woodland appearing in a fragmented arrangement (Figure 5-1a). Similarly, transitions to water as the largest patch show up at the edges of lakes and the ocean, as seen in Figures 5-1c, d, and e, although water was not considered in the process of randomization (Table 4-1).

In each experiment, RAMS was run 11 times, each with different amounts of classification error ranging from zero to 50 percent. The effect of classification accuracy on simulated regional climate is then examined by comparing the behavior of simulated precipitation within this range of accuracies to determine patterns. Then, patterns of behavior are compared across experiments to investigate the impacts of model parameterizations.

5.5 Results

5.5.1 Basic experiment

In the basic experiment, RAMS was run with the KF convection scheme and without interior nudging. As in Chapter 4, the performance of RAMS over

March – May 2003 is first assessed by comparing the RAMS-simulated accumulated precipitation to the TRMM observations.

Figure 5-2 shows both the simulated accumulated rainfall from RAMS and the observed accumulated rainfall from TRMM. As discussed in Section 4.6.2 of Chapter 4 for the whole year of 2003, RAMS underestimated precipitation in some areas, especially near the left and right boundaries, which may be due to the effect of boundary nudging. But it captured some major features, such as over the Congo forest. The spatial distribution of simulated precipitation is fairly similar to the TRMM observations, especially considering that no attempt was made to "tune" model parameters and that our configuration of RAMS has a lower spatial resolution (50 km vs. about 27 km). In Figure 5-3, precipitation is compared over time. Domain-averaged daily precipitation is normalized to 1 for the sake of comparison for the study area. The correlation coefficient is 0.336 for the whole time period, and 0.438 when the spin-up time of the first 20 days is omitted. Fidelity to observation improved over time, and the cessation of the "long rains" (day 77) is well-replicated.

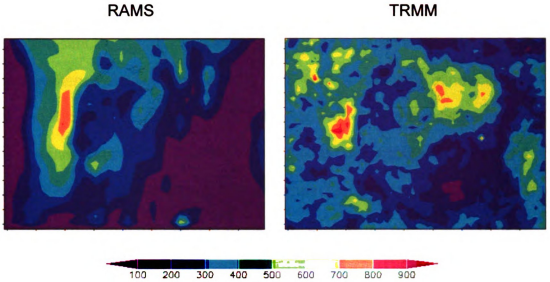


Figure 5-2. Spatial comparison of simulated accumulated precipitation (mm) in RAMS and that from TRMM for March – May, 2003

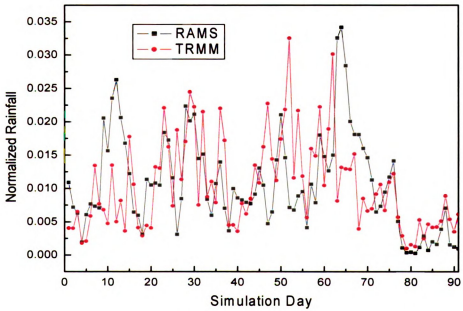


Figure 5-3. Temporal comparison of simulated cumulative precipitation in RAMS and from TRMM

The differences in precipitation between the simulation without land cover errors and simulations with errors (5%, 10% ... 45%, 50%) were examined. For the convenience of discussion, let R00, R05, R10 ... R45, R50 denote these 11 runs and R05-R00, R10-R00 ... R45-R00, R50-R00 denote the differences between runs. R10-R00, R30-R00, and R50-R00 are presented in Figure 5-4. If classification accuracy does not have any impact on simulated precipitation, then these differences are expected to be close to zero. However, as illustrated in Figure 5-4, precipitation differences are not minute. The impact on precipitation increases as classification accuracy worsens. It is also noticeable that most of the largest differences occur in the Lake Victoria area, even though errors are scattered across the whole domain (Figure 5-1).

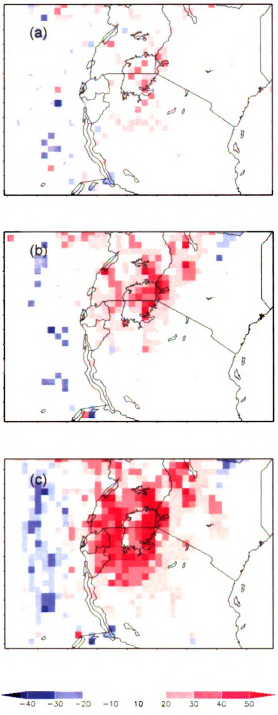


Figure 5-4. Differences of accumulated precipitation (mm) between simulation without land cover error and simulations with 10% error (R10-R00, a), 30% error (R30-R00, b), and 50% error (R50-R00, c).

Although a full investigation of this is beyond the scope of this dissertation, it is likely that the general spatial pattern of changes in precipitation shown in Figure 5-4 is due to the mechanism stated by Charney et al. (1977) and followed by other researchers (Lofgren 1995, Xue 1997, Wang et al. 2004). In this mechanism, a change in land surface parameters (e.g. albedo, vegetation fractional cover) alters the energy budget of the coupled surface-atmosphere system. Particularly at low latitudes, reduced heating of the atmosphere, resulting from increased surface albedo, leads to a relative sinking motion and reduced precipitation, while decreased surface albedo and increased atmospheric heating have the opposite effect.

In the domain considered here, the unperturbed GLC2000 land cover has nearly solid evergreen broadleaf forest in the western part of the domain, the class with the lowest surface albedo (Table 4-1). Thus, the insertion of random errors into this land cover class will necessarily increase the surface albedo, resulting in reduced precipitation in this area. Conversely, the region surrounding Lake Victoria is initially dominated by the crop/mixed farming class (in the GLC2000 dataset), which has the highest surface albedo of any of the classes used in the random replacement process, so the imposition of random errors reduces the surface albedo in this region. Combined with the ready access to water evaporating from Lake Victoria itself, this can lead to an increase in precipitation. As in Lofgren (1995), the heating of the atmosphere near Lake Victoria due to reduced surface albedo and the changes of other parameters is

likely compounded by the release of the latent heat of condensation associated with the increased precipitation.

Three measures were utilized to depict the precipitation differences between runs (Figure 5-4). The first measure is the maximum absolute difference (both positive and negative), which highlights only one hot spot. It represents the largest possible difference caused by land cover errors, but it does not give information on the overall differences. The other two measures used are the mean absolute difference and the standard deviation calculated over the whole domain. They characterize the overall magnitude and variation of the difference. The maximum absolute differences for R10-R00, R30-R00, and R50-R00 were 30.6mm, 56.7mm, and 84.4mm. The mean absolute differences were 4.6mm, 6.7mm, and 10.5mm while the standard deviations were 6.7mm, 9.5mm, and 14.7mm. The three measures all indicate an increase in precipitation difference as land cover accuracy decreases.

These three measures can evaluate precipitation differences against a range of classification errors (5% to 50%). For the basic experiment, the black lines in Figure 5-5 show the behavior of precipitation difference for this range of classification errors by illustrating the maximum and mean absolute differences and the standard deviation. In Figure 5-5a, the maximum absolute difference from the basic experiment increases from 34.8mm for 5% error to 84.4mm for 50% error. In Figure 5-5b, the mean absolute difference increases from 5.5mm to

10.5mm. The standard deviation increases from 7.8mm to 14.7mm (Figure 5-5c). From these three plots, it is evident that precipitation differences increase with an increase in land cover errors in RAMS. Importantly, when the errors are less than 20%, the plots are relatively flat, and when errors are larger than 20%, the differences increase sharply. This indicates that a classification error of less than 20% has little effect on the simulated precipitation in this particular experiment. The accuracy target of 85% commonly specified in the land cover production community can meet the requirements of regional climate modeling. If the land cover accuracy is less than 80%, however, its effect on climate simulation and propagation of uncertainty should be examined.

In the basic experiment results shown in Figure 5-5, the level area below 20% errors has non-zero differences. This is especially obvious for 5% error level. Adding this small amount of classification error causes some notable precipitation differences, which might be due to random noise. Above the 20% threshold, the signal rises above the noise. It is noticeable that there is a slight leveling off of the differences above the 40% level, which might be due to a saturation effect of classification errors.

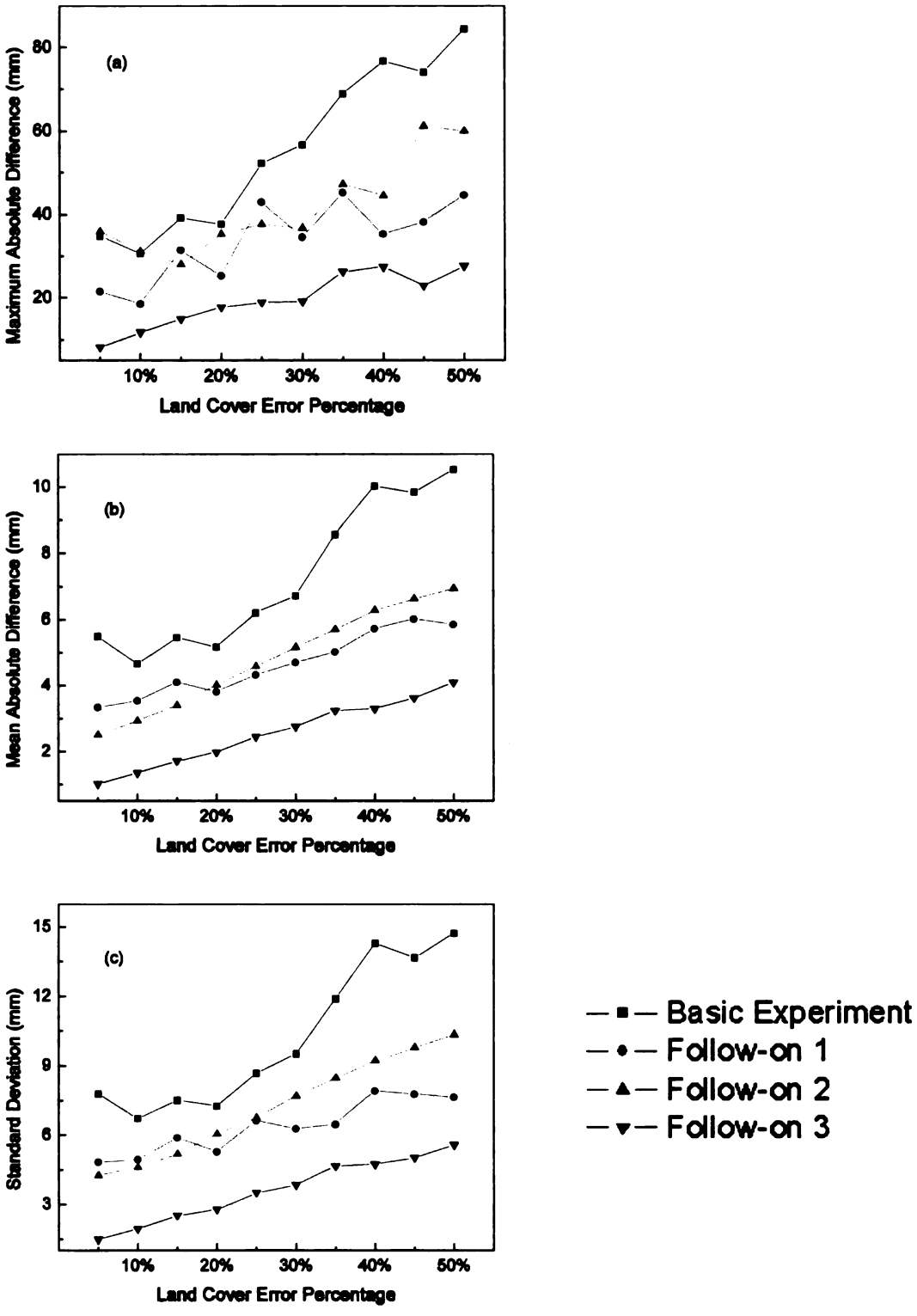


Figure 5-5. Maximum absolute differences (a), mean absolute differences (b) and standard deviations (c) from the basic experiment (Kain-Fritsch), follow-on 1 (Kuo), follow-on 2 (Kain-Fritsch with interior nudging), and follow-on 3 (Kuo with interior nudging).

5.5.2 Follow-on experiments

To test the effect of model parameterizations on atmospheric response to land cover accuracy, internal nudging and a different convection scheme are investigated. A different convection scheme may dramatically change the surface energy and moisture budget and, hence, surface feedback to the atmosphere. The KF convection scheme, which was used in the basic experiment, is known to produce more precipitation than the Kuo scheme, especially in areas of steep terrain (Castro et al. 2002, 2005). Nudging is used to relax the model solution towards the input reanalysis data continuously at each time step by adding artificial tendency terms (based on the difference between the two states) to the prognostic equations. With interior nudging, the surface boundary conditions tend to have weaker control on the vertical motion and distribution of precipitation, compared to no interior nudging (Weaver et al. 2002, Castro et al. 2005). Therefore, both convection scheme and interior nudging may influence the effect of land cover accuracy on simulated precipitation in RAMS. Other model aspects can also modify the influence of surface forcing on simulated precipitation; however, nudging and convection schemes are often used for such evaluations (e.g., Weaver et al. 2002, Castro et al. 2005).

In the basic experiment described previously, the KF convection scheme is used with no interior nudging, which allows the model to have a stronger response to surface boundary forcing. In follow-on experiment 1, the Kuo scheme was used with no interior nudging. In follow-on experiment 2, the KF

scheme was used with interior nudging applied and in the follow-on experiment 3, the Kuo scheme was used together with interior nudging. When interior nudging was used, the time scale was set to one day, which is larger than that specified in the RAMS User Guide (Castro et al. 2005). In each of these three follow-on experiments, RAMS was run 11 times, each with a different amount of classification error ranging from zero to 50%, similar to the basic experiment.

In Figure 5-6, accumulated precipitation is presented for the basic experiment and the three follow-on experiments, all with no classification error. As expected, the Kuo scheme produces much less precipitation over the whole domain. The major peak over the Congo forest, which is seen in experiments with the KF scheme and in satellite observations (Figure 5-2), is not shown clearly in experiments with the Kuo scheme. There is not much difference between these two schemes over dry areas, where both schemes tend to underestimate the precipitation. The interior nudging seems to have little effect on the accumulated precipitation.

In each follow-on experiment, precipitation differences between simulations with and without classification errors were investigated by examining the maximum and mean absolute differences and the standard deviation as in the basic experiment. In Figure 5-5, the behaviors of these three measures against a range of classification errors are presented for both the basic experiment and the three follow-on experiments. Precipitation in the follow-on

experiments is much less sensitive to classification errors, especially when the Kuo scheme was used. With both convection schemes, interior nudging tends to reduce this sensitivity. Interestingly, standard deviation plots for interior nudging are very close to straight lines. This may be due to the effect of interior nudging reducing the strength of small-scale variability, which has also been reported in other studies (e.g., Weaver et al. 2002).

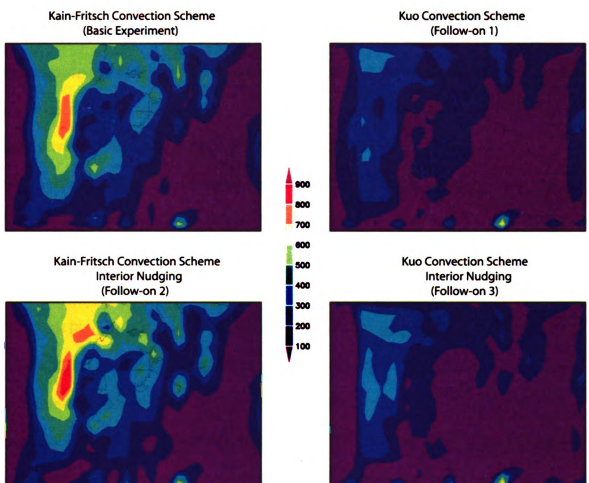


Figure 5-6. Simulated accumulated precipitation (mm) with different model parameterizations. Results from runs without land cover error are presented.

5.6 Discussion

In RAMS, each land cover type is represented by a suite of biophysical variables: albedo, leaf area index (LAI), fractional vegetation cover, etc. These biophysical variables determine the energy and moisture exchanges between the land surface and overlying atmosphere. Thus, the effect of land cover classification accuracy on simulated precipitation is ultimately controlled by the changes in the biophysical variables. Therefore, the effect of classification accuracy relies on how the surface scheme (LEAF-2 in this study) defines these biophysical variables for each type. As the biophysical parameters of different land cover types become more differentiated, the effect observed in previous sections will be more pronounced. In the hypothetical case where all land covers have exactly the same biophysical characteristics, classification accuracy will not have any effect on simulated precipitation.

In Figures 4-6 and 4-7, the default RAMS LAI and VFC are compared to the MODIS LAI and VFC respectively. It is evident that the LAI and VFC in RAMS are unrealistically uniform over most of the domain, with several regions poorly represented. Other biophysical parameters in RAMS version 4.4, such as albedo, may also have this characteristic of being overly homogeneous since they are defined by simple mathematical functions. Therefore, it is reasonable to expect that the impact of classification accuracy on simulated precipitation might be even greater than described in this study.

As shown in the previous sections, a land cover accuracy lower than 80% can substantially affect simulated precipitation, especially when the surface has a greater control of the atmosphere. This effect becomes stronger as the accuracy decreases. Although an 85% accuracy target has already been recommended for land cover production, in reality this target is rarely obtained (Trodd 1995). For example, the IGBP DISCover Land Cover product, a land cover layer from GLCC with global coverage, has an overall accuracy of 66.9%, which is considerably lower than the specified target (Scepan 1999). The default land cover dataset in RAMS (OGE) does not come with an accuracy estimate. Global accuracy for the newly developed MODIS Land Cover product (V003) is stated to be approximately 70 – 80% (http://geography.bu.edu/landcover/userguidelc/consistent.htm). When these global land cover products are used for a specific region, such as East Africa, the accuracy levels can be much lower than the global accuracy. Therefore, caution is needed when using global land cover products at regional to local levels. It should also be mentioned that global land cover products are usually developed for land cover identification or other general use. In order to be used in SVAT schemes in climate models, they usually need to be cross-referenced with biophysical parameters, which can add additional uncertainties.

These global products, however, have their advantages. The importance of classification accuracy is well recognized by producers. Quantitative

evaluation is therefore conducted to provide guidelines for users. Many land cover datasets employed in climate modeling studies do not come with accuracy information. This is especially true for historically reconstructed and future projected land cover datasets that are often employed to examine the impact of human activities on climate. Historical land covers are usually derived from existing maps and other indirect evidence, while future projected land covers are often developed from spatial models that simulate how changes in land use are likely to affect land cover. There are simply not many options for accuracy assessment of these types of land cover datasets. Historical and future land cover datasets are usually used to simulate time periods that are decades or even centuries long, much longer than the three months simulated in this study. The impact of land cover accuracy may well increase over these longer time frames. Uncertainties in those input land cover datasets may cause great uncertainties in the output in climate models.

There are aspects in this work that can be further explored. One is the strategy that was used to simulate classification errors. It is assumed that classification errors occur randomly over space. In reality, they are more likely to occur in areas with greater land surface heterogeneity and not in homogeneous landscapes such as the Congo forest. Also, in this study the original land cover types are replaced by random types from the five predominant types without considering the biophysical similarities between types. Land cover types with

similar physical appearances or similar spectral features in satellite images are more likely to be misclassified.

A second aspect relates to the configuration of RAMS. Factors such as horizontal grid spacing and multiple nested grids may influence the effect observed in this paper. The mosaic method of accounting for sub-grid variability in land cover does not take into account certain factors. Notably, latent, sensible, and radiative heat fluxes will be dependent on the land cover and on the characteristics of the air in the planetary boundary layer. The boundary-layer atmospheric characteristics are likely to be spatially correlated with the land cover, but the mosaic approach does not account for this and thus will miss the non-linear effects on fluxes. The situation is further complicated by sub-grid mesoscale circulations that can be forced by land cover heterogeneity (e.g., Weaver and Avissar 2001). These factors can be sensitive to the scale over which land cover is altered in addition to the model grid spacing. Further investigations considering these factors are needed.

5.7 Conclusions

Human activities have substantially modified the Earth's surface in the past and will continue to do so in the future. The impact of human activities such as land cover change on the regional and global climate can be studied using

climate modeling techniques. Land cover datasets, often derived from remote sensing images, are widely used in land surface schemes in climate models to describe the physical surface conditions. These datasets are not perfect, and their value is a function of classification accuracy. In the land cover production process, quantitative accuracy assessment has almost become a required procedure. However, the uncertainty arising from the accuracy of land cover classification and its impact on simulated climate have largely been ignored in climate modeling research.

In this paper, the Regional Atmospheric Modeling System was utilized to study the impact of land cover accuracy on simulated precipitation for the East Africa region. Classification errors were simulated as random alterations to the land cover dataset used in this study - GLC2000. The behavior of simulated accumulated precipitation over a three month period was then examined over a range of land cover errors (zero to 50%). It was found that, when the surface boundary had greater control on the overlaying atmosphere, land cover accuracy under 80% had a strong effect on simulated precipitation. As land cover accuracy worsened, this effect became even stronger. This effect on simulated precipitation was shown to be moderated by model parameterizations such as convection schemes and interior nudging, which affect the strength of control that the surface exerts on the atmosphere. When the Kuo convection scheme was used, RAMS severely underestimated the precipitation over the entire domain, and the land cover accuracy had little effect on simulated precipitation. With

interior nudging activated, the effect of land cover accuracy also decreased, even though the overall magnitude of precipitation was affected only slightly.

Based on the results of this study, it can be concluded that land cover datasets can meet general needs in climate modeling research if the commonly recommended 85% accuracy target is obtained. In reality, however, this is usually not the case. The reliability of land cover datasets needs to be examined in climate modeling research, especially those using historically reconstructed or future projected land covers for long-term simulations.

Chapter 6

Conclusions and Future Research

The accurate description of the spatial heterogeneity and seasonal variation of the landscape has been recognized as a key component in meaningful regional and global climate modeling. Land cover datasets and related biophysical variables are often used to represent the land surface in the state-of-art climate models e.g., the Regional Atmospheric Modeling System (RAMS) in this study. However, the land cover dataset in RAMS (OGE) was developed more than a decade ago on the basis of satellite images with limited capabilities, and the biophysical characteristics including leaf area index (LAI) and vegetation fractional cover (VFC) are depicted by a lookup table and simple mathematical equations.

This dissertation is devoted to improve the regional climate modeling in East Africa by better representing the land surface using various newly developed remote sensing products. Several new land cover products, together with the default OGE, are evaluated by a new statistical measure designed specifically for climate modeling purposes. The land cover with the best quality is selected to replace OGE land cover. LAI and VFC imageries from MODIS are incorporated directly in RAMS. The examination of the impacts of the improved land surface becomes possible by utilizing the land surface temperature (LST) and precipitation from remote sensing retrievals. For regions with very sparse

weather station observations such as East Africa, the focus of this study, model evaluations are extremely challenging tasks. Furthermore, the uncertainty from the land cover classification and its propagation in climate models, which have not been fully addressed, are investigated in this study. The threshold of classification accuracy is then identified, which can be used as a guideline for the land cover mapping in the remote sensing community as well as for various climate modeling studies, when historically reconstructed and future predicted land cover datasets are employed.

Assuming better classification has more consistent biophysical characteristics within each class, the new statistical measure Q developed for land cover classification accuracy assessment in this study aggregates the within-class LAI variation across the domain and over any time period. Unlike traditional evaluation methods which are based on ground-truthing data, Q utilizes newly developed MODIS LAI products and thus there is no sampling error. Another major advantage of Q is that it considers the consistency of classification over any specified time period when the seasonal variation of vegetation is important. More importantly, Q is designed on the basis of the mosaic approach of representing land surface heterogeneity in climate models. It is therefore more suitable for evaluating land cover datasets for climate modeling studies.

In terms of Q, GLC2000 ranks higher than MODIS land cover, OGE, and LEAF at three different spatial resolutions: 30 km, 50 km, and 100 km. Particularly, GLC2000 is significantly better than LEAF at both 30 km and 50 km. 50 km is the RAMS grid spacing in this study. Although, the newly developed MODIS land cover is better than OGE and LEAF at all three levels, the difference is not statistically significant. This evaluation suggests that the default land cover in RAMS needs to be replaced by GLC2000. In addition, this study finds as quadrate size decreases, the difference between land cover products tend to decrease.

The built-in LAI and VFC are compared spatially and temporally with corresponding MODIS products. Spatially, the default LAI and VFC are too homogeneous to differentiate distinct land surface types across the domain. Temporally, simple cosine functions have been demonstrated to be extremely unrealistic. For low-latitude regions, the prescribed LAI and VFC are almost invariant as time evolves in a year. To better represent the land surface conditions, spatially and temporally explicit MODIS LAI and VFC imageries are directly ingested in RAMS. The OGE land cover is replaced by GLC2000 as well. With insufficient weather station data available, the MODIS LST and TRMM precipitation data are used to evaluate the model performance and examine the impacts of improved land surface conditions. This study finds that the improved land surface improves the LST characteristics spatially and temporally but has little influence on simulated precipitation. Specifically, the bimodal feature of the

LST seasonal variation, which is completely missed in the default configurations, is fully captured when MODIS LAI and VFC are incorporated. Regarding to the spatial characteristics, the ITCZ-related seasonal migration of LST in the eastern domain has been greatly enhanced. Both MODIS Terra and Aqua LST are used for the first time to evaluate the surface impact on diurnal LST characteristics. This study finds that diurnal LST cycles in the second half of the year are slightly improved due to the new land surface representation. GLC2000 alone, however, is not sufficient to modify the seasonality of LST but is able to change the overall magnitude by about one degree.

In order to truly improve the land surface conditions in climate modeling, the uncertainty originating from the land surface and its propagation in climate models need to be better understood. The impact of the land cover classification accuracy on regional climate simulations remains a relatively unanswered question. In this study, a range of random classification errors are simulated and added to the baseline GLC2000. Differences in accumulated precipitation from March to May 20003 are examined. This study finds that the overall classification accuracy under 80% has a strong effect on simulated precipitation, especially when the land surface has a greater control on the overlying atmosphere. This effect becomes much stronger as the classification continues to worsen. In addition, model parameterizations, e.g. convection schemes and interior nudging, can influence this effect by mitigating the driving forcing of the land surface. This study suggests that land cover datasets developed from remote sensing

imageries can meet the requirement of regional climate modeling if the overall classification accuracy reaches 85%. In reality, however, this is usually not the case especially when the land cover datasets in climate modeling studies are historically reconstructed or future projected. In order to make convincing simulations, the reliability of the land cover dataset and the uncertainty propagation thus need to be fully investigated.

In the future, extended research is necessary to address the following issues related to this study. Firstly, the property of the Q statistic needs to be further studied. So far, its theoretical distribution and ranges under different circumstances are still not quite clear. Without knowing these properties, it is hard to tell whether a specific Q value is absolutely high or low. For land cover classification, a specific range of Q as a quality control is necessary if Q is applied in the land cover production community as an accuracy guideline. Secondly, other newly developed biophysical products such as MODIS albedo, have the potential to be incorporated into RAMS until the look-up table approach is completely discarded. Before this is achieved, the land surface is still partially improved. Of course, remote sensing techniques extracting other biophysical variables such as rooting depth need to be further developed. Thirdly, for the land cover classification accuracy, random classification errors are simulated in this study. In reality, misclassification may more likely occur between closely similar types such as shrubland and woodland not shrubland and forest. More realistic misclassification simulation approach thus needs to be implemented.

APPENDIX

Acronym List

ANOVA	One Way Analysis of Variance
AVHRR	Advanced Very High-Resolution Radiometer
BATS	Biosphere-Atmosphere Transfer Scheme
CCM	Community Climate Model
CLIP	Climate-Land Interaction Project
EVI	Enhanced Vegetation Index
GCM	Global Circulation Model
GLC2000	Global Land Cover 2000
IGBP	International Geosphere-Biosphere Programme
ITCZ	Intertropical Convergence Zone
LAI	Leaf Area Index
LEAF	Land Ecosystem-Atmosphere Feedback
LST	Land Surface Temperature
LUCC	Land Use/Cover Change
MCP	Multiple Comparison Procedure
MODIS	Moderate Resolution Imaging Spectroradiometer
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
OGE	Olson Global Ecosystem dataset
RAMS	Regional Atmospheric Modeling System
SST	Sea Surface Temperature
SVAT	Soil-Vegetation-Atmosphere Transfer
TRMM	Tropical Rainfall Measuring Mission
UTC	Coordinated Universal Time
VFC	Vegetation Fractional Cover

BIBLIOGRAPHY

Avissar, R. and Pielke, R. A. (1989). A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. Monthly Weather Review, 117, 2113-2136.

Bastidas, L. A., Gupta, H. V., Sorooshian, S., Shuttleworth, W. J., Yang, Z. L. (1999). Sensitivity analysis of a land surface scheme using multicriteria methods. Journal of Geophysical Research, 104, 19481-19490.

Becker, F. (1987). The impact of spectral emissivity on the measurement of land surface temperature from a satellite. International Journal of Remote Sensing, 8, 1509-1522.

Betts, R. A. (2000). Offset of the potential carbon sink from boreal forestation by decreases in surface albedo. Nature, 408, 187-190.

Bonan, G. B., Pollard, D., Thompson, S. L. (1992). Effects of boreal forest vegetation on global climate. Nature, 359, 716-718.

Bonan, G. B. (1996). A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Techincal description and user's guide. NCAR Tech. Note NCAR/TN-417 + STR, 150 pp.

Brown, J. F., Loveland, T. R., Ohlen, D. O., Donald, O., Zhu, Z. (1999). The global land-cover characteristics database: the user's perspective. Photogrammetric Engineering and Remote Sensing, 65, 1069-1074.

Carlson, T. N. and Ripley, D. A. (1997). On the relationship between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing of Environment, 62, 241-252.

Castro, C. L., Cheng, W. Y. Y., Beltran, A. B., Marshall, C. H., Jr., Pielke, R. A., Cotton, W. R. (2002). The incorporation of the Kain-Fritsch cumulus parameterization scheme in RAMS with a terrain-adjusted trigger function. The Fifth RAMS Users and Related Applications Workshop, Santorini, Greece.

Castro, C. L., Pielke, R. A., Leoncini, G. (2005). Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modelding System (RAMS). Journal of Geophysical Research, 110, D05108, doi:10.1029/2004JD004721.

Charney, J., Quirk, W. J., Chow, S-H, Kornfield, J. (1977). A comparative study of the effects of albedo change on drought in semi arid regions. Journal of the Atmospheric Sciences, 34, 1366-1385.

Chase, T. N., Pielke, R. A., Kittel, T. G. F., Nemani, R., Running, S. W. (1996). Sensitivity of a general circulation model to global changes in leaf area index. Journal of Geophysical Research, 101, 7393-7408.

Chase, T. N., Pielke, R. A., Kittel, T. G. F., Baron, J. S., Stohlgren, T. J. (1998). Potential impacts of historical land cover changes on global climate in northern winter. Climate Dynamics, 16, 93-105.

Chase, T. N., Peilke, R. A., Kittel, T. G. F., Nemani, R., Running, S. W. (2000). Simulated impacts of historical land cover changes on global climate in northern winter. Climate Dynamics, 16, 93-105.

Chen, C. and Cotton, W. R. (1983). A one-dimensional simulation of the stratocumulus-capped mixed layer. Boundary-Layer Meteorology, 15, 289-321.

Choudhury, B. J., Ahmed, N. U., Idso, S. B., Reginato, R. J., Daughtry, C. S. T. (1994). Relations between evaporation coefficients and vegetation indices studied by model simulations. Remote sensing of Environment, 50, 1-17.

Cihlar, J. (1997). GCOS/GTOS plan for terrestrial climate related observations. Version 2. GCOS Report. Geneva: World Meteorological Organization.

Cihlar, J. (2000). Land cover mapping of large areas from satellites: status and research priorities. International Journal of Remote Sensing, 21, 1093-1114.

Cohen, W. B., Maiersperger, T. K., Yang, Z., Gower, S. T., Turner, D. P., Ritts, W. D., Berterretche, M., Running, S. W. (2003). Comparison of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: a quality assessment of 2000/2001 provisional MODIS products. Remote Sensing of Environment, 88, 233-255.

Congalton, R. G. (1991). A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing of Environment, 37, 35-46.

Cotton, W. R., et al. (2003). RAMS 2001: Current status and future directions. Meteorology and Atmospheric Physics, 82, 5-29.

Cunnington, W. M. and Rowntree, P. R. (1986). Simulations of the Saharan atmosphere-dependence on moisture and albedo. Quarterly Journal of the Royal Meteorological Society, 112, 971-999.

Curran, P. J. (1983). Multispectral remote sensing for the estimation of green leaf area index. Philosophical Transactions of the Royal Society (London), 309, 257-270.

DeFries, R., Hansen, M., Steininger, M., Dubayah, R., Sohlberg, R., Townshend, J. (1997). Subpixel forest cover in Central Africa from multisensor, multitemporal data. Remote Sensing of Environment, 54, 209-222.

DeFries, R. S., Bounoua, L., Collatz, G. J. (2002). Human modification of the landscape and surface climate in the next fifty years. Global Change Biology, 8, 438-458.

Deardorff, J. W. (1978). Efficient prediction of ground surface temperature and moisture, with inclusion of layer of vegetation. Journal of Geophysical Research, 83, 1889-1903.

Di Gregorio, A. and Jansen, L. J. M. (2000). Land cover classification system (LCCS). Rome: F.A.O.

Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J., Wilson, M. F. (1986). Biosphere-atmosphere transfer scheme for the NCAR community climate model. NCAR Technical Note NCAR/TN-275+STR, Boulder, CO, USA.

Dickinson, R. E., Henderson-Sellers, A., Rosenzweig, C., Sellers, P. J. (1991). Evapotranspiration models with canopy resistance for use in climate models, a review. Agricultural and Forest Meteorology, 54, 373-388.

Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor, C. F. Ropelewski (2007). Validation of satellite rainfall products over East Africa's complex topography. International Journal of Remote Sensing, 28, 1503-1526.

Ducharne, A. and K. Laval (2000). Influence of the realistic description of soil water-holding capacity on the global water cycle in a GCM. Journal of Climate, 13, 4393-4413.

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

Foody, G. M. (2002). Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80, 185-201.

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

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao,

F., Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment, 83, 287-302.

Fritz, S., Bartholome, E., Belward, A., Hartley, A., Stibig, H.-J., Eva, H., et al. (2003). Harmonization, mosaicking, and production of the Global Land Cover 2000 database. Ispra, Italy: Joint Research Center (JRC).

Giorgi, F. and Mearns, L. O. (1999). Introduction to special section: regional climate modeling revisited. Journal of Geophysical Research, 104, 6335-6352.

Giri, C., Zhu, Z. and Read, B. (2005). A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. Remote Sensing of Environment, 94, 123-132.

Gutman, G. and Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. International Journal of Remote Sensing, 19, 1533-1543.

Hahmann, A. N. and Dickinson, R. E. (1997). RCCM2-BATS model over tropical south America: applications to tropical deforestation. Journal of Climate, 10, 1944-1964.

Henderson-Sellers, A., Dickinson, R. E., Durbidge, T. B., Kennedy, P. J., McGuffe, K., Pitman, A. J. (1993). Tropical deforestation: modeling local- to regional-scale climate change. Journal of Geophysical Research, 98, 7289-7315.

Houghton, R. A., Hackler, J. L., Lawrence, K. T. (1999). The U.S. carbon budget: contributions from land-use change. Science, 285, 574-578.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83, 195-213.

IGBP (1990). The international Geosphere-Biosphere Programme: a study of global change – the initial core projects. IGBP Report 12, Stockholm, Sweden.

IPCC (2001). Climate change 2001: the scientific basis. Cambridge University Press.

Jacob, D. and Podzun, R. (1997). Sensitivity studies with the regional climate model REMO. Meteorology and Atmospheric Physics, 63, 119-129.

Jansen, L. J. M. and Gregorio, A. D. (2003). Land-use data collection using the "land cover classification system": results from a case study in Kenya. Land Use Policy, 20, 131-148.

Jin, M., Dickinson, R. E., Vogelmann, A. M. (1997). A comparison of CCM2-BATS skin temperature and surface-air temperature with satellite and surface observations. Journal of Climate, 10, 1505-1524.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P., Morisette, J. T. (2002). An overview of MODIS Land data processing and product status. Remote Sensing of Environment, 83, 3-15.

Kain, J. S. and Fristch, J. M. (1993). Convective parameterization for mesoscale models: the Kain-Fristch scheme, in The Representation of Cumulus Convection in Numerical Models, Meteorological Monographs, 24, pp. 165-170, American Meteorological Society, Boston, Mass.

Kerr, Y. H., Lagouarde, J. P., Imbernon, J. (1992). Accurate land surface temperature retrieval from AVHRR data with use of an improved split-window algorithm. Remote sensing of Environment, 41, 197-209.

Kleidon, A. and Heimann, M. (2000). Assessing the role of deep rooted vegetation in the climate system with model simulations: Mechanism, comparison to observations and implications for Amazonian deforestation. Climate Dynamics, 16, 183-199.

Kummerow, C., Simpson, J., Thiele, O., Barnes, W., Chang, A. T. C., Stocker, E., Adler, R. F., Hou, A., Kakar, R., Wentz, F., Ashcroft, P., Kozu, T., Hong, Y., Okamoto, K., Iguchi, T., Kuroiwa, H., Im, E., Haddad, Z., Huffman, G., Ferrier, B., Olson, W. S., Zipser, E., Smith, E. A., Wilheit, T. T., North, G., Krishnamurti, T., Nakamura, K. (2000). The status of the tropical rainfall measuring mission (TRMM) after two yeas in orbit. Journal of Applied Meteorology, 39, 1965-1982.

Kuo, H. L. (1974). Further studies of the parameterization of the influence of cumulus convection on large-scale flow. Journal of Atmospheric Sciences, 31, 1232-1240.

Lambin, E. F. et al. (1999). Land-use and land-cover change (LUCC): implementation strategy. IGBP Report, 48, IGBP, Stockholm, Sweden.

Lark, R. M. (1995). Components of accuracy of maps with special reference to discriminant analysis on remote sensing data. International Journal of Remote Sensing, 16, 1461-1480.

Latifovic, R. and Olthof, I. (2004). Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. Remote Sensing of Environment, 90, 153-165.

Laval, K. (1986). General circulation model experiments with surface albedo changes. Climate Change, 9, 91-102.

Lean, J. and Rowntree, P. R. (1997). Understanding the sensitivity of a GCM simulation of Amazonian deforestation to the specification of vegetation and soil characteristics. Journal of Climate, 10, 1216-1235.

Lee, T. J. (1992). The impact of vegetation on the atmospheric boundary layer and convective storms. Atmosphere Science Paper No. 509, Colorado State University, CO, USA.

Lofgren, B. M. (1995). Surface albedo-climate feedback simulated using two-way coupling. Journal of Climate, 8, 2543-2562.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., Meachant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing, 21, 1303-1330.

Lu, L. and Shuttleworth, W. J. (2002). Incorporating NDVI-derived LAI into the climate version of RAMS and its impact on regional climate. Journal of Hydrometeorology, 3, 347-362.

Lynch, A. H., McIlwaine, S., Beringer, J., Bonan, G. B. (2001). An investigation of the sensitivity of a land surface model to climate change using a reduced form model. Climate Dynamics, 17, 643-652.

Mahlman, J. D. (1997). Uncertainties in projections of human-caused climate warming. Science, 278, 1416-1417.

Mathews, E. (1983). Global vegetation and land use: new high resolution data bases for climate studies. Journal of Climate and Applied Meteorology, 22, 474-487.

Mayaux, P., Bartholome, E., Fritz, S., Belward, A. (2004). A new land-cover map of Africa for the year 2000. Journal of Biogeography, 31, 861-877.

McGuffie, M. and Henderson-Sellers, A. (2001). Forty yeas of numerical climate modeling. International Journal of Climatology, 21, 1067-1109.

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

Molders, N., Strasser, U., Schneider, K., Mauser, W., Raabe, A. (1997). A sensitivity study on the initialization of surface characteristics in meso-β / γ modeling using digitized vs. satellite derived land-use data. Contributions to Atmospheric Physics, 70, 173-187.

Myneni, R. B. and Williams, D. L. (1994). On the relationship between FPAR and NDVI. Remote Sensing of Environment, 49, 200-211.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., Running, S. W. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sensing of Environment, 83, 214-231.

Nemani, R. R., Running, S. W., Pielke, R. A., Chase, T. N. (1996). Global vegetation cover changes from coarse resolution satellite data. Journal of Geophysical Research, 101, 7157-7162.

Nicholson, S. E., Tucker, C. J., Ba, M. B. (1998). Desertification, drought, and surface vegetation: an example from the west African Sahel. Bulletin of the American Meteorological Society, 79, 815-829.

Oleson, K. W. and Bonan, G. B. (2000). The effects of remotely sensed plant functional type and leaf area index on simulations of boreal forest surface fluxes by the NCAR land surface model. Journal of Hydrometeorology, 1, 431-446.

Olson, J. S., Watts, J. A., Allison, L. J. (1983). Carbon in live vegetation of major world ecosystems. Oak Ridge, TN: Oak Ridge National Laboratory.

Olson, J. S. (1994a). Global ecosystem framework-definitions. Internal report, USGS EROS Data Center, Sioux Falls, SD.

Olson, J. S. (1994b). Global ecosystem framework-translation strategy. Internal report, USGS EROS Data Center, Sioux Falls, SD.

Palmer, T. N. (2000). Predicting uncertainty in forecasts of weather and climate. Reports on Progress in Physics, 63, 71-116.

Pauwels, V. R. N. and Wood E. F. (2000). The importance of classification differences and spatial resolution of land cover data in the uncertainty in model results over boreal ecosystems. Journal of Hydrometeorology, 1, 255-266.

Pielke, R. A., et al. (1992). A comprehensive meteorological modeling system – RAMS. Meteorology and Atmospheric Physics, 49, 69-91.

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

Pielke, R. A. (2001). Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. Review of Geophysics, 39, 151-177.

Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. International Journal of Climatology, 23, 479-510.

Price, J. C. (1984). Land surface temperature measurements from the split window channels of the NOAA-7 AVHRR. Journal of Geophysical Research, 79, 5039-5044.

Price, J. C. (1992). Estimating vegetation amount from visible and near-infrared reflectances. Remote sensing of Environment, 41, 29-34.

Qin, Z. and Karnieli, A. (1999). Progress in the remote sensing of land surface temperature and ground emissivity using NOAA-AVHRR data. International Journal of Remote Sensing, 20, 2367-2393.

Reynolds, R. W. and Smith, T. M. (1994). Improved global sea surface temperature analyses using optimum interpolation. Journal of Climate, 7, 929-948.

Richardson, L. F. (1922). Weather prediction by numerical process. Cambridge: Cambridge University Press.

Running, S. W., Loveland, T. R., Pierce, L. L., Nemani, R. R., Hunt, E. R., Jr. (1995). A remote sensing based vegetation classification logic for global land cover analysis. Remote Sensing of Environment, 51, 39-48.

Running, S., Baldocchi, D., Turner, D., Gower, S., Bakwin, P., Hibbard, K. (1999). A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS data. Remote Sensing of Environment, 70, 108-127.

Salomonson, V. V., Barnes, W. L., Maymon, P. W., Montgomery, H. E., Ostrow, H. (1989). MODIS: advanced facility instruments for studies of the earth as a system. IEEE Transactions on Geoscience and Remote Sensing, 27, 145-153.

Scepan, J. (1999). Thematic validation of high-resolution global land-cover data sets. Photogrammetic Engineering and Remote Sensing, 65, 1051-1060.

Scheffe, H. (1959). The analysis of variance. New York: John Wiley & Sons.

Sellers, P. J., Mintz, Y., Sud, Y. C., Dalcher, A. (1986). A simple biosphere model (SiB) for use within general circulation models. Journal of the Atmospheric Sciences, 43, 505-531.

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D., Bounoua, L. (1996a). A revised land surface

parameterization (SiB2) for atmospheric GCMs. Part 1: model formulation. Journal of Climate, 9, 676-705.

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

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

Shabanov, N. V., Wang, Y., Buermann, W., Dong, J., Hoffman, S., Smith, G. R., Tian, Y., Hnyazikhin, Y., Myneni, R. B. (2003). Effect of foliage spatial heterogeneity in the MODIS LAI and FPAR algorithm over broadleaf forests. Remote Sensing of Environment, 85, 410-423.

Simpson, J., Adler, R. F. and North, G. R. (1988). A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. Bulletin of the American Meteorological Society, 69, 278-295.

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. Part I: the basic experiment. Monthly Weather Review, 91, 99-164.

Stapleton, J. H. (1995). Linear statistical models. New York: John Wiley & Sons.

Sud, Y. C. and Fennessy, M. J. (1982). A study of the influence of surface albedo on July circulation in semi-arid regions using the GLAS GCM. Journal of Climatology, 22, 105-125.

Sun, J. and Mahrt, L. (1995). Determination of surface fluxes from the surface radiation temperature. Journal of Atmospheric Sciences, 52, 1096-1106.

Taylor, C., Lambin, E. F., Stephenne, N., Harding, R., Essery, R. (2002). The influence of land-use change on climate in the Sahel. Journal of Climate, 15, 3615-3629.

Thomlinson, J. R., Bolstad, P. V., Cohen, W. B. (1999). Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products. Remote Sensing of Environment, 70, 16-28.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanow, N. V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B., Hame, T., Zndersson, K., Ozdogan, M., Knyazikhin, Y., Myeni, R. B. (2002). Multiscale analysis and

validation of the MODIS LAI product I. Uncertainty assessment. Remote Sensing of Environment, 83, 414-430.

Torbick, N., Lusch, D., Qi, J., Moore, N., Olson, J. and Ge, J. (2006). Developing land use/land cover parameterization for climate-land modeling in East Africa. International Journal of Remote Sensing, 27, 4227-4244.

Tremback, C. J. and Kessler, R. (1985). A surface temperature and moisture parameterization for use in mesoscale models. Seventh Conference on Numerical Weather Prediction, Montreal, Canada, American Meteorological Society, 355-358.

Trodd, N. M. (1995). Uncertainty in land cover mapping for modeling land cover change. Proceedings of RSS95 Remote Sensing in Action, Remote Sensing Society, Nottingham, UK.

Valor, E. and Caselles, V. (1996). Mapping land surface emissivity from NDVI: Application to European, African, and South American areas. Remote Sensing of Environment, 57, 167-184.

Verbyla, D. L. (2005). Assessment of the MODIS leaf area index product (MOD15) in Alaska. International Journal of Remote Sensing, 26, 1277-1284.

Vidale, P. L., Luthi, D., Frei, C., Seneviratne, S. I., Schar, C. (2003). Predictability and uncertainty in a regional climate model. Journal of Geophysical Research, 108, doi:10.1029/2002JD0028 10.

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

Wan, Z. and Li, Z.-L. (1997). A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. IEEE Transactions on Geoscience and Remote Sensing, 35, 980-996.

Wan, Z., Zhang, Y., Zhang, Q., Li, Z. (2002). Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. Remote Sensing of Environment, 83, 163-180.

Wan, Z., Zhang, Y., Zhang, Q. (2004). Quality assessment and validation of the MODIS global land surface temperature. International Journal of Remote Sensing, 25, 261-274.

Wang, J. R. (1992). Active and passive microwave measurements of soil moisture in FIFE. Journal of Geophysical Research, 97, 18979-18985.

Wang, G., Eltahir, E. A. B., Foley, J. A., Pollard, D., Levis, S. (2004). Decadal variability of rainfall in the Sahel: results from the coupled GENESIS-IBIS atmosphere-biosphere model. Climate Dynamics, 22, 625-637.

Wang, Y., Woodcock, C. E., Buermann, W., Stenberg, P., Voipio, P., Smolander, H., Hame, T., Tian, Y., Hu, J., Knyazikhin, Y., Myneni, R. B. (2004). Evaluation of the MODIS LAI algorithm at a coniferous forest site in Finland. Remote Sensing of Environment, 91, 114-127.

Washington, W. M. and Parkinson, C. L. (2005). An introduction to three-dimensional climate modeling (second edition). University Science Books: Mill Valley, Ca.

Watson, R. T., Noble, I. R., Bolin, B., Bavindranath, N. H., Verardo, D. J., Dokken, D. J. (2000). Land use, land use change, and forestry. IPCC Special Report, Intergovernmental Panel on Climate Change, Genewa, Switzerland.

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

Weaver, C. P., Roy, S. B., Avissar, R. (2002). Sensitivity of simulated mesoscale atmospheric circulations resulting from heterogeneity to aspects of model configuration. Journal of Geophysical Research, 107(D20), 8041, doi:10.1029/2001JD000376.

Wittich, K.-P. (1997). Some simple relationships between land-surface emissivity, greenness and the plant cover fraction for use in satellite remote sensing. International Journal of Biometeorology, 41, 58-64.

Xue, Y. (1997). Biosphere feedback on regional climate in tropical North Africa. Quarterly Journal of the Royal Meteorological Society, 123, 1483-1515.

Yang, Z-L. (2003). Modeling land surface processes in short-term weather and climate studies. Theory and Modeling of Atmospheric Variability, edited by Zu, X., Li, X., Cai, M., Zhou, S., Zhu, Y., Jin, F.-F., Zou, X., Zhang, M., World Scientific Series on Meteorology of East Asia, World Scientific, New Jersey, 288-313.

Zeng, X., Dai, J., Dickinson, R. E., Shaikh, M. (1998). The role of root distribution for climate simulation over land. Geophysical Research Letter, 25, 4533-4536.

Zeng, X., Dickinson, R. E., Walker, A., Shaikh, M., DeFries, R. S., Qi, J. (2000). Derivation and evaluation of global 1-km fractional vegetation cover data for land modeling. Journal of Applied Meteorology, 39, 826-839.

Zhao, M., Pitman, A. J., Chase, T. N. (2001a). Climatic effects of land cover change at different carbon dioxide levels. Climate Research, 17, 1-18.

Zhao, M., Pitman, A. J., Chase, T. N. (2001b). The impact of land cover change on the atmospheric circulation. Climate Dynamics, 17, 467-477.

Zhuang, X., Engel, B. A., Xiong, X., Johannsen, C. J. (1995). Analysis of classification results of remotely sensed data and evaluation of classification algorithms. Photogrammetric Engineering and Remote Sensing, 61, 427-433.