

San Jose State University SJSU ScholarWorks

Master's Theses

Master's Theses and Graduate Research

Summer 2013

Evaluation and Improvements of the Offline CLM4 Using ARM Data

Terrence J. Mullens San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd theses

Recommended Citation

Mullens, Terrence J., "Evaluation and Improvements of the Offline CLM4 Using ARM Data" (2013). *Master's Theses.* 4356. DOI: https://doi.org/10.31979/etd.xu6t-vx9c https://scholarworks.sjsu.edu/etd_theses/4356

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

EVALUATION AND IMPROVEMENTS OF THE OFFLINE CLM4 USING ARM DATA

A Thesis

Presented To

The Faculty of the Department of Meteorology and Climate Science

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Masters of Science

by

Terrence J. Mullens

August 2013

© 2013

Terrence J. Mullens

ALL RIGHTS RESERVED

The Designated Thesis Committee Approves the Thesis Titled

EVALUATION AND IMPROVEMENTS OF THE OFFLINE CLM4 USING ARM DATA

by

Terrence J. Mullens

APPROVED FOR THE DEPARTMENT OF METEOROLOGY AND CLIMATE SCIENCE

August 2013

Dr. Menglin S. JinDepartment of Meteorology and Climate ScienceDr. Martin J. LeachDepartment of Meteorology and Climate ScienceDr. Samuel LevisThe National Center for Atmospheric Research

ABSTRACT

EVALUATION AND IMPROVEMENTS OF THE OFFLINE CLM4 USING ARM DATA

by Terrence J. Mullens

Hourly ground observations for year 2004 from the Atmospheric Radiation Measurement (ARM) program of the Department of Energy were used to examine the surface and subsurface energy simulations of the Community Land Model version 4 (CLM4). The 2 m air temperature, wind speed, solar radiation, downward longwave radiation, and precipitation observed by the ARM project were used to force the offline CLM4, and the ARM land surface and soil observations including skin temperature $(T_{\rm skin})$, soil temperature and moisture, and sensible, latent, and ground heat fluxes were used to evaluate the model outputs. The default and ARM-forced CLM4 runs for 2004 were compared to assess the improvements to the model for hourly, daily, and seasonal timescales. The root mean square error and the Pearson correlation coefficient show that the ARM-forced offline CLM4 leads to improved accuracy in surface and soil energy fluxes in comparison with the default offline CLM4. Nevertheless, a warm bias of 2°C to 3°C was assessed on T_{skin} in summer due to warm maximum temperatures and in winter due to warm minimum temperatures. To improve CLM4 T_{skin} simulations, a proposed vegetation emissivity parameterization was evaluated locally and globally using both ARM and Moderate Resolution Imaging Spectroradiometer remote-sensing observations. This new algorithm results in cooling and an improvement of 0.17 K for the ARM site. Global evaluation revealed improvement in areas of intermediate canopy density.

ACKNOWLEDGMENTS

I'd like to thank my thesis committee: Dr. Menglin Jin, Dr. Martin Leach, and Dr. Samuel Levis. Dr. Jin has encouraged me and pushed me to my true potential, always offering guidance and insight into my work. Dr. Leach has been invaluable during my coursework at SJSU and in the writing process of this work. Dr. Levis gave substantial amounts of his time and energy during my visit to NCAR in 2012, offering assistance and suggestions, which substantially enhanced this work.

I also thank the other faculty members and friends I've encountered, who have fostered a positive environment during my time at SJSU. Dr. Clements, Dr. Bridger, Dr. Chiao, and Dr. Cordero all have provided feedback, encouragement, opportunities, and knowledge to my studies here. My friends Henry Bartholomew, Colin McKellar, Allison Charland, and Angela Reside, just to name a few, have provided friendship and excellent memories, which have made my time here something I will look back on fondly. I'm grateful for my fiancée, Jamie, who has supported me throughout my time here. I'd also like to thank my family, both blood and in Christ, for their love and support, with special acknowledgment to the people of Garden City Church, who have given me a solid spiritual community to be a part of during my time here.

Finally, I'd like to thank my Lord and Savior, Jesus Christ, who has stayed along my side throughout my time here. May all my works, here and beyond, be my gift to him.

This work was funded by the NSF Large-Scale Dynamics Program (award number 0855480, PI-Jin) and the NCAR Advanced Study Program.

List of Figures	viii
List of Tables	xi
1. Introduction and Background	1
a. Introduction	1
b. Land Surface Processes in Climate Models	5
c. History of CLM Validations	
2. Data, Model, and Experimental Design	11
a. Model Description	11
b. ARM Data Sets	13
c. Model Simulations and Evaluation	18
3. ARM-Forced Improvements and Evaluation	
a. Daily Calculated Surface Energy Evaluation	
b. Daily Soil Temperature and Moisture Evaluation	
c. Seasonal Surface Energy Evaluation	
d. Hourly Temperature Evaluation and Soil Temperature Lag	
e. Discussion	
4. Conclusions	
a. Future Work	
b. Final Remarks	50
Peteroneos	50

TABLE OF CONTENTS

Appendix A: List of Acronyms	. 59
	<i>c</i> 1
Appendix B: Improvement to Vegetation Canopy Emissivity	. 61

LIST OF FIGURES

1.	Location of the ARM SGP site in Lamont, Oklahoma. Adapted from the	
	Department of Energy ARM project (www.arm.gov) 15	
2.	Site layout at the ARM-Lamont Central Facility. Adapted from Stokes and	
	Schwartz (1994)	
3.	Time series of monthly ARM and MODIS Terra T_{skin} values over Lamont,	
	Oklahoma, for 2004	
4.	Daily T_{skin} for (a) CLM4 versus ARM T_{skin} from the ARM facility in Lamont,	
	Oklahoma, for 2004. (b) Daily T_{skin} sensitivity to ARM-forcing. (c) Daily T_{skin}	
	scatterplot between CLM4 ARM-forced T_{skin} and observations	
5.	Comparisons among daily SH flux for control (CLM4), CLM4 ARM-forced, and	
	ARM observations at the Lamont, Oklahoma, for 2004	
6.	Comparisons among daily LE flux for control (CLM4), CLM4 ARM-forced, and	
	ARM observations at the Lamont, Oklahoma, for 2004	
7.	Comparisons among daily ground heat flux for control (CLM4), CLM4 ARM-	
	forced, and ARM observations at the Lamont, Oklahoma, for 2004 25	
8.	Surface albedo calculations for the default CLM4, ARM-forced CLM4, and ARM	
	observations at the Lamont, Oklahoma, ARM facility for 2004	
9.	Calculated ARM-forced CLM4 fractional snow cover and observed MODIS-Terra	
	fractional snow cover observations for the region covering Lamont, Oklahoma, for	
	1 January to 30 April 2004	

- 12. Offline CLM4-simulated ARM-observed volumetric soil moisture for (a) CLM layer 3 (0.05 m) and (b) CLM layer 5 (0.25 m), for the year 2004. The control run is forced by the default atmospheric data from Qian et al. (2006), and the ARM-forced run (CLM4 ARM-forced) is forced using ARM atmospheric observations... 32
- 14. Scatterplot of daily CLM4 ARM-forced versus ARM-observed G flux values, errors, and correlations for (a) winter, (b) spring, (c) summer, and (d) fall for 2004.

- 2004 and (b) July 2004 for Lamont, Oklahoma...... 40

18.	8. Daily time series for averaged CLM4 ARM-forced modeled air temperature,	
	$T_{\rm skin}$, and soil temperatures for (a) January 2004 and (b) July 2004 for Lamont,	
	Oklahoma. Both cycles begin at Midnight CST and end at Midnight CST each	
	day	42
B1.	Time series of satellite observed ELAI and ESAI used in CLM4 over	
	Lamont, Oklahoma and time series of default CLM4 canopy emissivity compared	
	to values using the proposed algorithm and observation	70
B2.	Daily T_{skin} simulations for ARM-forced CLM4 using the default equation versus	
	ARM-forced CLM4 using proposed equation	72
B3.	Daily T_{skin} simulations for ARM-forced CLM4 using default emissivity equation	
	versus ARM-Forced CLM4 using proposed equation	74
B4.	Map of global ELAI+ESAI and global differences in RMSE calculations for	
	skin temperature between the default CLM4 and the CLM4 run using the	
	proposed equation	76

LIST OF TABLES

1. ARM observations used as forcing and evaluation data in this experiment	5
2. Table of CLM4 soil layer depths	12
B1. Proposed PFT-based emissivity values used in this experiment	67

1. Introduction and Background

a. Introduction

Reliable general circulation models (GCMs) are critical tools to predict, mitigate, and adapt to natural and human-induced climate changes (Bernstein et al. 2007; Williams et al. 2009; Brown et al. 2012). Because it is unfeasible to study the entire climate system through experimental methods, GCMs have become the primary tools for scientists to study climate change (Edwards 2011). These models use known physical laws, parameterizations, and governing equations to predict how components in the climate system will respond to anthropogenic changes. Originally, GCMs focused primarily on the atmosphere, so other components were represented by a set of observations and constants or were disregarded altogether (Dickinson et al.1995; Sellers et al. 1997). In the past two decades, however, GCMs have integrated the ocean, land surface, sea ice, urban centers, and carbon cycle to their models, with each being an individual subcomponent interacting with one another.

Land models are a critical component of GCMs because the energy, momentum, and moisture exchanges at the land surface are critical to local and global climate (Sellers et al. 1986; Dickinson 1995; Sellers et al. 1997; Nicholson 1998; Dickinson et al. 2006; Baklanov et al. 2011). Furthermore, land models simulate how the biosphere responds to climate change (Bonan et al. 2002a). Land models have developed from simple "bucket" methods (Manabe 1969) to a complex system that can model land and biogeophysical processes, such as surface energy fluxes and vegetation evolution on the land surface. For example, the Community Land Model (CLM; Bonan et al. 2002b; Dai et al. 2003; Oleson et al. 2010) managed at the National Center for Atmospheric Research (NCAR), models ecosystem evolution, vegetation dynamics, transport of moisture, and surface energy fluxes throughout the climate system. These physical processes are crucial to climate modeling because they partially determine the availability of energy and moisture in the atmosphere.

There is substantial uncertainty in the performance of GCMs. The accurate simulation of observed climate variability increases confidence in a model's performance (Brown et al. 2012). Similar to GCMs, land models contain uncertainties and thus require validation. The validation of past land models has led to the incorporation of better physical, ecological, and hydrological parameterizations in the models. Examples include the introduction of plant functional types (PFTs), the improvement of fractional snow cover using Advanced Very High Resolution Radiometer (AVHRR) data (Bonan et al. 2002a; Niu and Yang 2007), and improvements to land and vegetation cover parameters using the Moderate Resolution Imaging Spectroradiometer (MODIS) data (Lawrence and Chase 2007). In addition, single-point tower data sets, such as FLUXNET data (Baldocchi et al. 2001; Williams et al. 2009) and GAME-Tibet data (<u>http://monsoon.t.u-</u> tokyo.ac.jp/tibet), have been used to evaluate both localized and global simulations of land models (Stockli et al. 2008; Zeng et al. 2012). Furthermore, the strengths of various models have been identified through validation, allowing new models to include the strengths of previous models. For example, the CLM inherits the strengths of three separate land models: the Biosphere-Atmosphere Transfer Scheme (Dickinson et al.

1993), the NCAR Land Surface Model (LSM1 and LSM2; Bonan 1996; Oleson and Bonan 2000), and the Chinese Academy of Science Land Model (IAP 94; Dai and Zeng. 1997).

The validation of land surface models identifies weaknesses in a model. For example, validations have revealed that the simulation of latent heat flux is a weakness in CLM and in other climate models (Leuning et al. 2012; Lawrence et al. 2012). Furthermore, validation has revealed weaknesses in the simulations of fractional snow cover (Niu and Yang 2007; Swenson and Lawrence 2012), vegetation cover (Lawrence and Chase 2007), and sensible heat flux (Zeng et al. 2012). Knowledge of these weaknesses leads to modifications to improve land models. These modifications undergo the same validation process that the original model underwent to determine if the modification improves the simulation while maintaining the integrity of the rest of the model.

High-quality in situ data sets to force and to evaluate land models provide an accurate validation approach (Bonan 2008; Williams et al. 2009). Currently, typical land model evaluations use multiple data sets taken through different field campaigns (Bonan et al. 2002b; Dai et al. 2003; Niu and Yang 2007; and Zeng et al. 2012). These data sets are valuable in evaluating model performance. However, differing measurement techniques and resolutions of different campaign data sets can produce uncertainties when used in forcing and evaluating land models. Therefore, a set of observed data measured simultaneously at the same site would be appropriate for model evaluation.

The Atmospheric Radiation Measurement (ARM; Stokes and Schwartz 1994; Ackerman and Stokes 2003) project aims to provide various data sets for climate model evaluations. For example, the ARM data have been proven successful in validating model parameterizations, such as albedo (Yang et al. 2008), as well as surface energy fluxes and their relationships with cloud fraction (Qian et al. 2012). Although atmospheric measurements from ARM have been used to evaluate models, land surface observations have not been adequately used. For example, measurements of land surface temperature (LST) and soil fluxes have not yet been used to evaluate land models. Because similar field campaigns have succeeded in evaluating and improving land models, the ARM project, which has more instrumentation and staff at each site compared with similar campaigns, could potentially be a major contribution to land model validation.

This thesis focuses on evaluating the offline NCAR Community Land Model version 4 (CLM4; Oleson et al. 2010; Lawrence et al. 2011; Lawrence et al. 2012). The goal is to validate CLM4 using ARM observations that have been developed into forcing (ARM-forcing) and evaluation (ARM-evaluation) data sets (Table 1). An offline CLM4 run was forced using the default atmospheric data for 2004 (Qian et al. 2006) and the observed ARM-forcing data for the year 2004. Afterward, both runs were evaluated with the ARM-evaluation data set at hourly, daily, and monthly timescales. The results of this study highlight physical processes that need further improvement. Furthermore, the results suggest particular seasons and times of day that improvements should be focused on. Finally, an example of improvements to vegetation canopy emissivity (ε_v) made through this validation is given in Appendix B.

Forcing Data	Evaluation Data
2 m air temperature (<i>T</i>)	Upward shortwave radiation (S_{up})
Direct solar radiation (S_{dir})	Upward longwave radiation
	• Used to calculate $T_{\rm skin}$
Diffuse solar radiation (S_{dif})	Sensible heat flux
Relative humidity	Latent heat flux
Wind speed (<i>V</i>)	Soil temperature at 5 and 25 cm depth
Precipitation	Soil moisture at 5 and 25 cm depth

Table 1. ARM observations used as forcing and evaluation data in this experiment.

b. Land Surface Processes in Climate Models

The Earth's surface and atmosphere interact through various geophysical and chemical processes in the planetary boundary layer (Dickinson 1995; Sellers et al. 1997; Nicholson 1998; Dickinson 2000; Dai et al. 2003; Baklanov et al. 2011; Jin and Mullens 2012). Thus, an understanding of the structure, composition, and evolution of the land surface is crucial in climate modeling. Furthermore, Dirmeyer et al. (2012) suggest that global warming potentially increases the influence of the land surface on geophysical exchanges in the climate system, specifically between the atmosphere and the land surface; therefore, properly modeling these exchanges is necessary to predict future impacts. CLM4 models the land surface; the various exchanges of heat, moisture, and gas (such as CO₂); and the ecological elements between the land surface and the atmosphere,

such as dynamic vegetation and parameterizations of urban areas. As these processes are introduced into the modeled climate system, the accuracy of future climate prediction may increase substantially.

The surface energy balance (SEB) is an important relationship between the land surface and the atmosphere. Incoming shortwave radiation from the sun and downward longwave radiation emitted by the atmosphere are distributed in terms of heat and moisture fluxes, as shown in the following SEB equation:

$$\frac{S_{d}(1-\alpha)}{(A)} + \frac{L_{d}}{(B)} - \frac{\varepsilon \sigma T_{skin}^{4}}{(C)} - \frac{SH}{(D)} - \frac{LE}{(E)} - \frac{G}{(F)} = 0$$
(1)

- A. A percentage of the incoming solar shortwave radiation (S_d) is reflected off the land surface, based on the surface albedo (α), the ratio of reflected radiation to incoming radiation at the surface. The rest is absorbed at the surface.
- B. Some longwave radiation is emitted from clouds and the atmosphere, directed toward the Earth's surface (L_d).
- C. An amount of the energy absorbed by the land surface is emitted as longwave radiation. This is a function of the radiometric surface temperature, also known as skin temperature (T_{skin} ; Jin et al. 1997; Jin 2004; Jin and Dickinson 2010), and the emissivity (ε) of the land surface. This relationship is quantified through the Stefan-Boltzmann law.

D. Because the land surface and the atmosphere often have different temperatures, energy is exchanged between them to achieve thermodynamic equilibrium. This is called *sensible heat* (SH) flux, which is highest when the temperature difference between the land surface and the atmosphere is at its highest level. For example, in CLM4, the SH flux between the ground and the atmosphere is calculated as follows (Oleson et al. 2010, Eq. 5.61):

$$H = -\rho_{\rm atm} C_{\rm p} \frac{(\theta_{\rm atm} - T_{\rm G})}{r_{\rm ah}}, \qquad (2)$$

where ρ_{atm} and θ_{atm} are the density and the potential temperatures of the atmosphere, respectively; C_{p} is the specific heat of the air; T_{g} is the temperature of the ground; and r_{ah} is the aerodynamic resistance to SH transfer.

E. Some of the energy is involved in the phase changes of water, which either absorb or emit heat through evaporation or condensation near the land surface. This is called *latent heat* (LE) flux. In CLM4, this quantity is a product of λ, the LE of vaporization (or sublimation if air temperature is below freezing point), and the water vapor flux, calculated as follows (Oleson et al. 2010, Eq. 5.62):

$$E = -\rho_{\rm atm} \frac{\beta_{\rm soi}(q_{\rm atm} - q_{\rm g})}{r_{\rm aw}},$$
 (3)

where β_{soi} is an empirical function of soil water, q_{atm} is the specific humidity of the atmosphere, q_g is the specific humidity at the ground surface, and r_{aw} is the aerodynamic resistance to water vapor transfer. F. The remaining energy that is not reflected, emitted, exchanged between the land and the atmosphere or used in the phase changes of water is then absorbed into the ground. This is called the heat flux into the ground or, more commonly, the ground (G) flux.

All of the terms in the SEB equation are important and must be accurately simulated in GCMs. Atmospheric conditions (e.g., the presence of clouds, humidity, air temperature, and greenhouse gases) and land properties (e.g., snow cover, land use, soil moisture, and vegetation cover) affect the SEB. Therefore, these parameters substantially affect the transfer of energy and moisture between the land and the atmosphere and thus the Earth's climate (Wiscombe and Warren 1980; Anthes 1984; Dirmeyer 2000). All of these parameters are modeled in CLM4 and measured by the ARM project.

c. History of CLM Validations

CLM0, the first version of the model (Zeng et al. 2002), was compared with LSM1 and LSM2 using observational data from Willmott and Matsurra (2000) and Valadi hydrological data (Schlosser 1996). Later, Bonan et al. (2002b) demonstrated the improvements of CLM2 over LSM1 and LSM2. In particular, simulated snow water equivalent, 2 m air temperature (T_{air}), precipitation, and runoff were improved when CLM2 was coupled with NCAR's Community Climate Model (CCM3). In addition, Dai et al. (2003) used the same runs from Bonan et al. (2002b) but included additional offline

validations using the Valadi data together with data from the Anglo-Brazilian Amazonian Climate Observation Study (Gash et al. 1996). These studies demonstrated CLM's ability to simulate land surface properties in comparison with CLM's predecessors, even with the same offline forcing data. Lawrence and Chase (2007) used MODIS data to develop new land surface parameters (such as vegetation types) in the offline CLM3.0. Their study showed encouraging accuracy of the offline CLM and clear improvement gained by using observations. Roesch (2006), Niu and Yang (2007), and Wang and Zeng (2010) have used various satellite remote-sensing (such as AVHRR and MODIS) and surface observation data sets to evaluate snow cover and albedo and how their modifications to the algorithms improved the CLM's simulations. Jin and Liang (2006) used the MODIS and the National Center for Environmental Prediction (NCEP)-NCAR reanalysis data to improve bare soil emissivity. Qian et al. (2006) used a set of global and station data to validate the hydrological elements of the model and a new set of forcing data that is currently used. FLUXNET data have been used to evaluate single-point simulations of carbon, water, and energy fluxes (Baldocchi et al. 2001; Williams et al. 2009). MODIS data and snow depth observations have been used to improve modeled fractional snow cover (Swenson and Lawrence 2012). Furthermore, observational data have been assimilated into the model through the Data Assimilation Research Testbed (Anderson et al. 2009).

The CLM serves as the land surface component of NCAR's Community Earth Systems Model (CESM; Oleson et al. 2010), a major contributor to the 2007 Intergovernmental Panel on Climate Change (IPCC) fourth assessment report (Dickinson et al. 2006; Bernstein et al. 2007). The CLM underwent extensive validation (Bonan et al. 2002b; Dai et al. 2003; Qian et al. 2006) to demonstrate its performance. Because of CLM's role as a component of CESM, and thus the IPCC reports, improvements made to the model may lead to improvements in forecasting future climate change.

In this study, ARM data sets were used to evaluate CLM4, the current version of the model. The advantage of using ARM observations is that the forcing data are collected with the evaluation data (at the same site and same time). Consequently, a more accurate one-to-one validation can be performed, thus removing uncertainties that occur otherwise due to inconsistencies between data sets. In addition, ARM provides data in a finer spatial and temporal scale than many forcing and evaluation data sets, which allows for a more robust evaluation of land models.

Section 2 outlines the description of the offline CLM4 as well as the ARMforcing and ARM-evaluation data collected from the CO₂ flux site in Lamont, Oklahoma, and the experimental design and methods used for evaluating the model. Section 3 discusses the results of offline CLM4 default runs and runs with the ARM-forcing and discussion of these results. Final remarks are made in Section 4. Appendix B proposes a new scheme to improve vegetation canopy emissivity (ε_v).

2. Data, Model, and Experimental Design.

a. Model Description

CLM4 (Oleson et al. 2010; Lawrence et al. 2012), the latest released version of the CLM, models ecosystem, groundwater, and surface energy fluxes. Recent developments in CLM4 include improved software and computational performance, competition between PFTs for water in a single column, improved hydrological cycle, and improved vegetation dynamics. Furthermore, improvements were also made in vegetation burial by snow and simulated fractional snow cover. A more detailed history and description of the model is provided by Oleson et al. (2010). The current status of the model along with new developments and challenges are outlined by Lawrence et al. (2012).

CLM4 splits the Earth into individual grid cells and calculates parameters for each cell. The cell is further split into a heterogeneous nested subgrid for different land units (glacier, wetland, vegetated, lake, or urban), which are then split into soil-snow columns and broken up into various PFTs (Bonan et al. 2002a). In addition, the model simulates soil temperature and soil moisture at 10 subsurface soil layers and 5 layers of bedrock (Table 2). CLM4 can be run in two different modes: fully coupled to the CESM or Community Atmosphere Model (the atmospheric component of CESM) or in offline mode with prescribed atmospheric forcing. CLM needs to be "spun up" prior to the period that model simulations will be run so that the model will stabilize after a climatologically abnormal period, such as a drought (Yang et al. 1995). The CESM (and

thus CLM4) runs on a "no-leap" calendar because of the algorithm used to calculate incoming solar radiation at the top of the atmosphere (Oleson et al. 2010; Neale et al. 2010), thus leap days are not modeled.

Layer	Depth (m)
1	0-0.018
2	0.018-0.045
3	0.045-0.091
4	0.091-0.166
5	0.166-0.289
6	0.289-0.493
7	0.493-0.829
8	0.829-1.383
9	1.383-2.296
10	2.296-3.433
11–15	Bedrock

Table 2. Subsurface soil layers and subsequent depths in CLM4.

The default offline CLM4 uses the atmospheric forcing data of Qian et al. (2006), which are a blend of NCEP-NCAR reanalysis data variations and some observationsbased analysis. The spatial resolution of the default forcing data is T62 (~1.825°), and the temporal resolution is 3-hourly (0000, 0300, etc., UTC) for six parameters: precipitation

.

rates, T_{air} , specific humidity, wind speed, surface air pressure, and solar radiation. These data are then bilinear interpolated to hourly data for model simulations, which require more frequent output and consequentially input. The default atmospheric forcing data set spans from 1948 to 2004.

However, the default forcing has a few disadvantages. First, the spatial resolution is too coarse when the model is being validated by surface data measured over a point far smaller than 1.825° in latitude and longitude. Second, the temporal resolution (3 h) is also too coarse. Because significant variation in conditions can occur hourly, a fine temporal resolution would be desired.

Therefore, to truly understand how the model is performing when compared with observed land surface data over a given point, a set of complementary atmospheric forcing data over the same point and at a fine temporal resolution is necessary. Although ARM observations themselves contain uncertainties (such as instrument calibration, drift errors, etc.), they remove a majority of the uncertainties included by the forcing data, allowing for a much better evaluation of the model's calculations themselves.

b. ARM Data Sets

A set of observed atmospheric and land surface data are recorded by the ARM Southern Great Plains (SGP) CO_2 flux tower (Fischer 2005; Fischer et al. 2007). The CO_2 flux site at Lamont measures CO_2 , SH, and LE fluxes, as well as solar and terrestrial

radiation every hour, while friction velocity and Monin-Obukhov scale length are calculated. The data from the CO_2 flux site are useful as forcing data and for evaluating the output of models. However, the data contain a large amount of missing values and uncertainties. To provide a more accurate atmospheric forcing data set, a value-added refinement of the data has been developed at San José State University as ARM-forcing and ARM-evaluation data sets, respectively. The ARM data underwent a quality control process, removing bad data and replacing missing data through linear interpolation or substitution of previous data points if the missing points span over too long a period to accurately interpolate (Jin et al. 2013, submitted). Most cases of missing data were solved by the first method, with only few instances requiring the second substitution method. Although these processes allow for a more continuous data set, they also introduce additional uncertainties in the data because missing data are simply replaced with interpolated data and therefore might not properly represent what is indeed occurring at that time point. Additional uncertainties from the data include the drift and calibration errors, which are common among instrumentation-based measurements, and the accuracy of the instrumentation, which is outlined in the CO_2 flux handbook (Fischer 2005). An additional uncertainty is caused by the positioning of the sonic anemometer used at this site, which can lead to an underestimation of SH flux by as much as 10% (Frank et al. 2013). For this study, 2004 was randomly chosen for initial data development. This data set, however, is being further developed for validations on a shorter timescale (such as a day or a month) for other years.

The ARM SGP site is located in Lamont, Oklahoma (36.6° N, 97.5° W; Figure 1) at an elevation of 1030 ft (314 m) in an area of open pasture blocked off from the surrounding farmlands (Figure 2). The region of Lamont, Oklahoma, is a humid subtropical climate (Koppen Classification Cfa), with hot, wet summers and cold, drier winters. The average high temperature for July is 93° F (33.9° C), and the average low temperature for January is 22° F (-5.6° C). The average annual precipitation is 35 inches of rain and 12 inches of snow (approximately 1 inch of equivalent rain). The vegetation of the surrounding farmlands is typically winter wheat but varies by season and year and does not include the ARM SGP central facility itself (i.e., the vegetation on the site does not vary with farming). This may have a profound impact on land cover features in land models because the model could be using a different vegetation input compared with the site itself.



Figure 1. Location of the ARM SGP site in Lamont, Oklahoma. Adapted from the Department of Energy ARM project (www.arm.gov).



Figure 2. Site layout at the ARM–Lamont Central Facility. Adapted from Stokes and Schwartz (1994).

At the ARM SGP CO_2 flux site, there are two soil sensors taking half-hourly measurements in the ground: one at 5 cm (0.05 m) below the ground and one at 25 cm (0.25 m) below the ground. These depths are modeled in layers 3 and 5 of CLM4, respectively. The primary measurements taken by these instruments (and used in this study) are soil moisture and soil temperature. Although the quality of these data is acceptable for much of 2004, there exists a large amount of missing data for soil temperature in the second half of 2004. Furthermore, soil moisture readings during the same period vary unrealistically. For the purpose of this study, that those uncertain and missing data have been removed from the comparison to produce the most accurate results possible.

The primary goal of this study is to evaluate the daily, diurnal, and seasonal simulations of T_{skin} . To have the best perspective on the validity of ARM data in evaluating T_{skin} , it is useful to compare the observations with another observed data set. To do so, monthly LST data, taken from the MODIS instrument on board the Terra satellite, are used. These data are described by Wan (2008), and the uncertainties are discussed by Jin and Mullens (2012). Because MODIS only takes two measurements of a given location each day (at approximately 10:30 a.m. and 10:30 p.m. local time), the values are averaged into a single temperature value for each month. The comparison indicates that although both observations have the same seasonality pattern, MODIS LST observations are warmer by as much as 3 K for most months (Figure 3). One possible reason for this is cloud cover. Because MODIS LST is determined based on the measurement of upward longwave radiation, LST can only be measured on clear days,

which likely subjects the measurements to a warm bias because the land surface is heated by higher solar radiation levels.



Figure 3. Time series of monthly ARM and MODIS Terra T_{skin} values over Lamont, Oklahoma, for 2004.

c. Model Simulations and Evaluation

Two offline CLM4 runs were performed for the year 2004. The first was a control run (CLM4) using the default (Qian et al. 2006) atmospheric forcing data, with a 50-year spin-up, performed over the globe. The grid point containing Lamont, Oklahoma, was extracted from the output to be evaluated with surface observations. The second was a single-point offline run (CLM4 ARM-forced) replacing the default Qian atmospheric forcing data with the ARM-forcing data we have developed. The run was performed with a 50-year spin-up, the initial soil moisture in the model was replaced with observations, and the land cover and the PFT weight were set as 0.4 C3 grasses, 0.4 C4 grasses, and 0.2

bare soil to represent the winter wheat and grass cover at the site. The temporal resolution of the model output was daily, from 1 January 2004 to 31 December 2004, with additional hourly output for January and July. The output of this run was then evaluated in the same manner as the default run, and the evaluations were compared with one another. Because there is no model output data for the leap day (29 February 2004), this date was removed from ARM-forcing data and observations.

To evaluate the accuracy of CLM4's output when compared with ARMevaluation data, the root mean square error (RMSE) and the Pearson correlation coefficient (r) were used. These metrics are calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (M - O)^2}, \quad (4)$$

$$r = \frac{n \sum MO - \sum M \sum O}{\sqrt{(n \sum M^2 - [\sum M]^2)(n \sum O^2 - [\sum O]^2)}},$$
 (5)

where *n* is the number of observations, *M* is the modeled output, and *O* is the observation. These quantities are the classical methods for evaluating model output (Taylor 2001). RMSE quantifies the differences between the two models, whereas *r* quantifies pattern similarity between the two models (Taylor 2001). Although these are classical methods for evaluating model output, Williams et al. (2009) argue that nonrandom errors and bias in fluxes prevent these methods from being optimal and suggest additional evaluation methods, specifically quantifying patterns at different frequencies. However, because evaluating *T*_{skin} is the primary goal of this study, the use of less traditional approaches in the evaluation produces greater uncertainty than using the classical approach. It is however worth noting that the statistical approach in this study may produce uncertainty.

This analysis was performed at hourly, daily, seasonal, and annual timescales to assess how the model performs at different timescales. For the seasonal analysis, winter was defined as the months of December, January, and February; spring was defined as March, April, and May; summer was defined as June, July, and August; and fall was defined as September, October, and November. An hourly analysis was also performed for January and July to determine whether the errors change diurnally during the winter and summer months. Investigating the errors and correlations sheds light on how the model performs when forced with the default forcing and again when forced with the improved ARM-forcing data.

Modifications to CLM4 vegetation canopy emissivity (ε_v) are proposed and evaluated in Appendix B. The default canopy vegetation parameter was replaced with a parameter dependent on both the canopy density and the PFT structure of the land surface. The sensitivity experiment uses the same ARM-forcing data as discussed previously, and the output was compared with the CLM4 ARM-forced run. The emissivity changes were also applied globally, using MODIS data for global evaluation purposes.

3. ARM-forced Improvements and Evaluation

a. Daily Calculated Surface Energy Evaluation

The CLM4 ARM-forced run produces an RMSE of 2.43 K for T_{skin} compared with surface observations, whereas the CLM4 run produces an RMSE of 3.12 K for T_{skin} (Figure 4a). This results in an improvement of 0.68 K. Because T_{skin} is a function of radiative fluxes and T_{air} , it can be suggested that the majority of improvements in modeled T_{skin} are due to improved solar radiation and T_{air} forcing. Furthermore, the use of ARM-forcing data in the model did not result in uniform warming or cooling throughout the year. The forcing caused a decrease in simulated T_{skin} on some days but an increase on others, as seen in Figure 4b.

When compared with ARM observations, CLM4 ARM-forced T_{skin} generally follows daily variations. The correlation between modeled T_{skin} and observations for CLM4 ARM-forced T_{skin} is 0.989 (Figure 4c). Although the modeled T_{skin} correlates well with observations, there are periods where modeled T_{skin} values are consistently higher than observations, specifically in the summer months (approximately days 160–250).



Figure 4. (a) Daily T_{skin} for CLM4 versus ARM T_{skin} from the ARM facility in Lamont, Oklahoma, for 2004. (b) Daily T_{skin} sensitivity to ARM-forcing. (c) Daily T_{skin} scatterplot between CLM4 ARM-forced T_{skin} and observations.

Sensible heat flux does not improve, in general, when forced with ARM-forcing data. CLM4 ARM-forced SH has an RMSE of 32.02 Wm^{-2} compared with an RMSE of 31.13 Wm^{-2} for CLM4 (Figure 5). On many days, however, the forcing decreases the amount at which the model either overestimates or underestimates SH, such as between days 10 and 15. Furthermore, there are days where the forcing brings the modeled SH to more reasonable values, such as days 315-334, where the modeled SH shifts from negative to positive values because of the forcing data. However, these improvements are counteracted by increases in difference during other periods, with the most problematic being between days 40-70 and 140-200.



Figure 5. Comparisons among daily SH flux for control (CLM4), CLM4 ARM-forced, and ARM observations at the Lamont, Oklahoma, for 2004.

In addition, simulated LE fluxes do not improve with ARM-forcing in general. ARM-forced LE has an RMSE of 36.52 W m^{-2} compared with an RMSE of 31.34 W m^{-2} for CLM4 (Figure 6). However, the degradation in accuracy is not uniform. For example, the forcing improves LE during periods where the day-to-day variation is lowest, specifically during days 0–70 and 240–329, which is during winter, early spring, and late fall. The greatest degradation in accuracy occurs from days 130 to 230, which is during late spring, summer, and early fall. This suggests that the model struggles with simulating LE during warm periods where LE varies greatly from day-to-day, regardless of forcing.



Figure 6. Comparisons among daily LE flux for control (CLM4), CLM4 ARM-forced, and ARM observations at the Lamont, Oklahoma, for 2004.

Although neither SH nor LE simulations show improvement, CLM4 ARM-forced ground flux simulations show some improvement. CLM4 ARM-forced G flux has an RMSE of 18.22 W m⁻², whereas CLM4 G flux has an RMSE of 20.53 W m⁻² (Figure 7). Much of this improvement occurs at the beginning and end of 2004, where the overestimation of G flux is somewhat tempered in the CLM4 ARM-forced simulations. This is seen clearly near days 10 and 330, where sharp decreases in G flux are moderated
by the forcing data. In general, the model overestimates day-to-day variations in G flux during winter months and underestimates variations during the summer months.



Figure 7. Comparisons among daily ground heat flux for control (CLM4), CLM4 ARMforced, and ARM observations at the Lamont, Oklahoma, for 2004.

Surface albedo is evidently improved with improved forcing and land cover changes (Figure 8). In general, albedo is closer to observations in the CLM4 ARM-forced run, particularly in days 300–365, where CLM4 has consistently low albedo whereas CLM4 ARM-forced albedo is close to the observations. Nevertheless, evident deficiencies also occur: between days 40 and 280, observed albedo is always higher than 0.20, and both model runs have lower albedo, with the CLM4 run being lowest. In addition, during snow events at days 30 and 40 (Figure 8), observed albedo is higher than 0.60 whereas modeled albedo is at most the same as days without snow. This suggests that the model may simulate some snow coverage but cannot simulate snow albedo well.



Figure 8. Surface albedo calculations for the default CLM4, ARM-forced CLM4, and ARM observations at the Lamont, Oklahoma, ARM facility for 2004.

Two possible reasons for the underestimation of albedo on high albedo days are the short duration of the high albedo events and the poor simulation of snow cover. To determine which one is most likely, a comparison of CLM4 fractional snow cover to daily observed fractional snow cover is useful. Because ARM does not measure snow cover, the fractional snow cover product retrieved from the MODIS instrument flown on the Terra satellite is used as an alternative. Further information on the daily MODIS snow cover products is given by Hall and Riggs (2007). The comparison between calculated model fractional snow cover and MODIS (Figure 9) shows that CLM4 substantially underestimates the fractional snow cover for the region. This suggests that the primary reason why CLM4 does not reproduce the same spikes in albedo as the ARM observations is the lack of snow on the ground rather than the duration of the event. The current snow algorithm in CLM4 (from Niu and Yang 2007) determines fractional snow cover by the depth of snow, which is a function of the amount of snow that has fallen. Therefore, neither of these precipitation events would be able to produce substantial fractional snow cover values in CLM4.



Figure 9. Calculated ARM-forced CLM4 fractional snow cover and observed MODIS-Terra fractional snow cover observations for the region covering Lamont, Oklahoma, for 1 January to 30 April 2004.

This is further illustrated by the actual precipitation forcing (Figure 10). Both instances of substantially high albedo and observed snow cover from MODIS do indeed correspond to a precipitation event (approximately days 30 and 41). However, both events were relatively light, resulting in an underestimation of fractional snow coverage by the model. In addition to snow events, sudden decreases in observed albedo

correspond to precipitation events due to the darkening of soil. Examples of this include days 20, 35, and 61. In all three events, the decrease in albedo was not simulated by either model run. This suggests that even with improved precipitation forcing, the model simulates neither snow coverage nor darkening of soil accurately.



Figure 10. Daily observed precipitation taken at the Lamont, Oklahoma, ARM facility and used as the precipitation forcing in the ARM-forced CLM4 case for the first 120 days of 2004.

b. Daily Soil Temperature and Moisture Evaluation

The CLM4 ARM-forced soil temperature RMSE improves from 2.07 K to 1.74 K in layer 3 (0.05 m depth) and from 4.74 K to 4.1 K in layer 5 (0.25 m depth; Figure 11). The main problems with CLM4 ARM-forced in layer 3 are that modeled winter temperatures are consistently lower than observed temperatures, and the variations in modeled soil temperature are slightly sharper than variations seen in the observed temperatures. In layer 5, CLM4 ARM-forced has a substantial cold bias compared with observations. Furthermore, there is a lag between temperature changes in the observations and changes modeled in layer 5, indicating that simulated heat flux from the upper layers of soil is not reaching this layer efficiently. In addition, because layer 3 has a lower RMSE compared to layer 5, it is suggested that errors in modeled heat flux occur between layers 3 and 5.



Figure 11. Offline CLM4-simulated versus ARM-observed soil temperatures for (a) CLM layer 3 (0.05 m) and (b) CLM layer 5 (0.25 m) for the year 2004. Substantial amounts of ARM-observed data were missing in the second half of 2004.

Modeled volumetric soil moisture does not improve substantially using ARMforcing. In layer 3, the default RMSE values are 0.0463 and 0.0457 mm³ mm⁻³ for the CLM4 ARM-forced. In layer 5, the default RMSE is 0.0358 mm³ mm⁻³, and the CLM4 ARM-forced RMSE is 0.0315 mm³ mm⁻³. Soil moisture in layer 3 is substantially drier than observations (Figure 12a). The model, however, still simulates the same variations as the observations, especially when the model is forced with the ARM atmospheric observations, likely due to improvements in precipitation values for the ARM-forced CLM4 compared with the default CLM4. Nonetheless, although the model responds to precipitation events, the increase in soil moisture due to precipitation does not bring the simulated amounts to the same level as observations. This can be seen on days 14, 30, 60, and 175 in layer 3, where the simulated values do not peak as high as the observations. In addition, soil moisture in layer 5 is generally wetter than observations (Figure 12b). Similar to layer 3, many of the peaks in soil moisture are underestimated compared with observations, specifically on days 14, 30, and 60. Although the model responds to precipitation in the same manner as layer 3, moisture does not leave layer 5 efficiently. This suggests that the moisture fluxes may be too rapid in upper layers but not efficient in lower layers. There is the possibility that the inability for water to efficiently move through layer 5 might be compounded by layer 3 getting rid of its water too rapidly. A possible explanation for this is the type of soil used in this experiment. Although the forcing data and the PFT structure were modified to better reflect the ARM site, the soil type is still determined through a global data set, which may lead to an incorrect soil type.

Because soils transport both heat and water differently, depending on soil type, the use of an incorrect soil type may lead to an inaccurate simulation of soil moisture.



Figure 12. Offline CLM4 simulated versus ARM-observed volumetric soil moisture for (a) CLM layer 3 (0.05 m) and (b) CLM layer 5 (0.25 m) for the year 2004. The control run is forced by the default atmospheric data from Qian et al. (2006), and the ARM-forced run (CLM4 ARM-forced) is forced using ARM atmospheric observations.

c. Seasonal Evaluation

Model performance varies at different temporal scales. The accuracy of T_{skin} in the ARM-forced CLM4 has a strong seasonal variation. Winter (Figure 13a) and summer (Figure 13c) prove to have the highest errors, with RMSE values of 2.75 K and 2.90 K, respectively. Furthermore, winter and summer are the lowest correlated between modeled and observed values of the four seasons, with r = 0.95 for winter and 0.92 for summer. Modeled T_{skin} calculations for winter have a warm bias present on the colder days. Summer T_{skin} has a substantial warm bias that is worse on warmer days. Both spring (Figure 13b) and fall (Figure 13d) are substantially less erroneous, with RMSE values of 1.14 K for spring and 2.44 K for fall. In addition, both spring and fall have a correlation of r = 0.99. Furthermore, the primary errors for spring and fall occur due to an increasingly warm bias with higher temperatures. These results suggest that moderate temperatures are modeled best in CLM4, whereas extreme temperatures are more subject to a warm bias and poorer correlation with observations.



Figure 13. Scatterplot of daily CLM4 ARM-forced versus ARM-observed T_{skin} values, errors, and correlations for (a) winter, (b) spring, (c) summer, and (d) fall for 2004.

Similar to daily averaged output, CLM4 exhibits difficulty in accurately calculating ground flux seasonally (Figure 14). Winter and spring have the highest errors and lowest correlations (winter: RMSE = 24.99 W m⁻², r = 0.54; spring: RMSE = 17.36 W m⁻², r = 0.36), whereas summer and fall are better correlated (summer: RMSE = 17.73

W m⁻², r = 0.62; fall: RMSE = 14.02 W m⁻², r = 0.65). There are many instances where CLM4 ARM-forced G flux is significantly more negative than observations during winter months. This suggests that a substantial amount of the winter cold bias in layer 3 soil temperatures (as discussed earlier) may be due to overestimates of negative ground flux on certain days. With the exception of a few days during spring, this problem is not present in other seasons.



Figure 14. Scatterplot of daily CLM4 ARM-forced versus ARM-observed ground flux values, errors, and correlations for (a) winter, (b) spring, (c) summer, and (d) fall for 2004.

Sensible heat flux, although poorly correlated to the observations, exhibits a seasonal pattern similar to that of T_{skin} (Figure 15). Spring and fall are best modeled, with RMSE values of 31.2 and 20.72 W m⁻², respectively. In addition, both seasons have better correlations with observations (r = 0.44 for spring and r = 0.66 for fall), whereas winter and summer are poorly modeled (winter: RMSE = 34.4 W m⁻², <u>r</u> = 0.18; summer: RMSE = 47.52, r = 0.06). Because SH flux depends on the temperature of the ground (and thus T_{skin}), it can be suggested that much of the similarity to T_{skin} in seasonality is due to T_{skin} simulations. However, only fall has much resemblance to observations. In summer, the model typically overestimates SH. This suggests that the overestimation of T_{skin} during this period may result in the overestimation of SH.

LE flux (Figure 16) has a similar pattern to SH flux with spring and fall being best correlated (r = 0.64 for spring and r = 0.73 for fall). However, calculated errors are lowest in winter (16.16 W m⁻²) and highest in summer (91.37 W m⁻²). A possible explanation is the amount of incoming solar radiation has the same seasonality, yielding more energy distribution in the summer and less in the winter. The lower incoming solar radiation leads to less evaporation, which results in lower amounts of LE being absorbed by water during the evaporation process. Nevertheless, the end result is that LE variation is lowest in winter and highest in summer, which suggests that errors are also lowest in winter and highest in summer. Furthermore, neither winter nor summer has a high

correlation (r = 0.3 for winter and r = 0.35 for summer), suggesting that although winter has the lowest RMSE, spring and fall are still the best simulated.



Figure 15. Scatterplot of daily CLM4 ARM-forced versus ARM-observed SH flux values, errors, and correlations for (a) winter, (b) spring, (c) summer, and (d) fall for 2004.



Figure 16. Scatterplot of daily CLM4 ARM-forced versus ARM-observed LE flux values, errors, and correlations for (a) winter, (b) spring, (c) summer, and (d) fall for 2004.

d. Hourly Temperature Evaluation and Soil Temperature Lag

The diurnal variation of modeled T_{skin} reveals a better understanding of when the model is most problematic. For January (which is used to represent the winter season), a comparison with hourly observations (Figure 17a) shows that there is a warm bias in the daily minimum temperatures, which is likely the primary cause of the seasonal warm bias

(as seen in Figure 13a). The rest of the diurnal cycle is modeled more accurately when compared with the observations, especially when forced by the ARM-forcing data. CLM4-ARM-forced produces results that are remarkably close to the observations, even on days where there is little variation (such as between hours 380 and 430 of July).

For the month of July, there is a substantial warm bias during the warmest time of the day (Figure 17b). Although the forcing data is able to temper this by keeping maximum temperatures close to observations on cooler days (especially on days of little variation, such as between hours 180 and 200 of July), the model still substantially overestimates the temperatures during the warmest part of the day, regardless of the forcing data. There also exists, on many days, a smaller warm bias in the daily minimum temperatures as well as a short lag between when the ARM observations reach minimum and when the modeled T_{skin} values reach bottom (such as hours 400 and 590 of July).



Figure 17. Hourly CLM4 ARM-forced versus ARM-observed T_{skin} values for (a) January, 2004 and (b) July 2004 for Lamont, Oklahoma.

One final issue in diurnal variation is the time lag between the maximum T_{air} , T_{skin} , and soil temperature layers in the model. Ideally, T_{skin} should have the largest daily

variation and a slight lead in maximum temperature when compared with T_{air} . Soil layers with increasing depth should have an increasing lag of maximum temperature and decreasing diurnal amplitude compared with T_{skin} . Indeed, this is what is seen for modeled T_{skin} and soil temperatures for both January and July 2004 (Figure 18). Both months model a T_{skin} diurnal cycle that closely follows the T_{air} cycle. There is a lead of 2 h compared with T_{air} in both January and July. Both months model T_{skin} to be colder than T_{air} at night, and both months model T_{skin} to be warmer than T_{air} during the day. The time lag between maximum T_{skin} and maximum soil temperature does indeed increase which increasing soil depth. In addition, the amplitude of the daily variation indeed decreases with increasing soil depth in January, but the maximum soil temperature in layer 3 becomes greater than maximum T_{air} in July. There is little diurnal variation in layer 5, and the maximum occurs approximately 10 h after the maximum T_{skin} , whereas layer 3 has higher amplitude and a maximum of approximately 4 h after the maximum T_{skin} . This is further discussed by Jin et al. (2013, submitted).



Figure 18. Daily time series for averaged CLM4 ARM-forced modeled air temperature, $T_{\rm skin}$, and soil temperatures for (a) January 2004 and (b) July 2004 for Lamont, Oklahoma. Both cycles begin at Midnight CST and end at Midnight CST each day.

e. Discussion

These results validate T_{skin} , surface energy, and moisture parameters modeled in CLM4 and provide insight into areas of the model that can be improved. An attempt at improving the model is discussed in Appendix B. Although additional work is needed in assuring the robustness of the improvements (namely, applying the changes proposed to other single-point sites), they demonstrate the usefulness of ARM data as a tool for improving CLM4.

When forced with ARM data, CLM4-modeled T_{skin} better follows both the diurnal and the seasonal cycles observed by the ARM-evaluation data than the model forced by the default (Qian) forcing. This is specifically seen in the winter, spring, and fall months, whereas the summer months produce a substantial warm bias in maximum temperatures. Because most of the year experiences improvements when forced with improved forcing data, the lack of improvement in RMSE in summer T_{skin} suggests that other elements in modeling T_{skin} are likely responsible for such a warm bias. One possible explanation could be the substantial errors in LE flux, which may result in the addition of extra energy into the SEB and an increase in upward longwave radiation and thus in T_{skin} .

The surface energy fluxes are a known problem in CLM4 (Lawrence et al. 2011; Lawrence et al. 2012). Interestingly, the errors in SH, LE, and G fluxes when forced with ARM-forcing data suggest that forcing data are not the primary cause of errors in the simulation of SH, LE, and G fluxes. In addition, model performance varies seasonally. Both SH and LE fluxes prove to be the most accurate during the spring and fall months and most problematic in the summer months. This may suggest, at least partly, that the errors in these fluxes may be responsible for the problematic calculations in T_{skin} during the summer months.

Furthermore, the poor measurement and modeling of SH and LE fluxes have been shown to extend beyond CLM4 and even beyond modeling to the observations. Leuning et al. (2012) discuss the problems in accurately measuring SH and LE fluxes, specifically due to phase lags caused by the incorrect estimation of energy storage below the land surface. This leads to an unbalanced SEB budget and thus a remainder term. This is called the *energy imbalance problem* (Foken 2008). Therefore, it is indeed possible that the poor RMSE and *R* values calculated between the ARM data and the CLM4 output may be partially caused by the energy imbalance problem in the observations, not just by

the model itself. In addition, Frank et al. (2013) suggest that sonic anemometers that do not measure vertical wind orthogonally underestimate SH fluxes by as much as 10%. Indeed, the sonic anemometer used at the CO_2 flux site does not measure the vertical wind orthogonally, making the observations subject to this error. Furthermore, because the simulations in this study use a single-point data set, horizontal wind effects are not represented in the calculation of SH.

Although modeled G is still relatively inaccurate, it is better correlated to observed G year-round than SH or LE. In addition, the RMSE does not vary greatly from season to season. Ultimately, these results suggest a need for improvement in the modeling of all three heat fluxes. Such improvements might even make issues related to the accurate modeling of T_{skin} less significant because they may be at the root of T_{skin} 's problems.

An additional surface parameter that vastly affects the distribution of surface energy is surface albedo. When compared with ARM observations, the model has consistently low albedo year-round. This problem is even worse during the presence of snow cover, which was poorly simulated in the model. The lack of modeled snow cover presents a challenge in the proper calculation of surface energy on snowy days. In addition, the model struggled to respond to precipitation events, namely, the decrease in albedo caused by the darkening of the soil from rainfall. Another possible issue with the calculation of albedo is the soil type. Because CLM4 uses a single soil type for each grid cell, the soil heterogeneity within the grid might not be properly represented. Future work

could include creating a mosaic of soil types that more accurately represents the ARM facility. Furthermore, the land cover at the site is in fact winter wheat, but the model currently does not have winter wheat as a land cover. This may also lead to unrealistic simulations.

Soil temperature and soil moisture also have deficiencies in the model. Although too much ARM data are missing from the summer months to allow for a robust seasonal comparison as was conducted for the surface, the day-to-day comparison is still helpful in evaluating the soil in CLM4. Although modeled daily variation in soil temperature in layer 3 is relatively accurate (and further improved with ARM-forcing), the soil temperature in layer 5 was consistently too cold. One of the possible reasons may be the movement of heat from the surface layer to the lower layers. Nevertheless, the accuracy of soil temperature in layer 3 suggests that the movement of heat is accurately calculated in the first several layers of the model most of the year. However, the calculation of heat flux below (at least) layer 3 may be problematic. Because CLM4 output defaults to providing the heat flux data at the surface and between layers 1 and 2, it is not possible to determine that a low-biased flux calculation near layer 5 is the cause for the cold bias in that layer. In addition, the lag in day-to-day temperature changes in layer 5 (clearly seen in Figure 11b) also suggests that heat is not getting to this layer as quickly as it should be. Again, this problem is not seen in layer 3, which suggests that the problem creating this is likely manifesting somewhere between layers 3 and 5.

The diurnal comparison of modeled T_{skin} with ARM observations suggests that the main problems in the accurate calculation of a daily T_{skin} value are a bias in the calculation of maximum and minimum temperature values, especially in the summer months. In January, the CLM4 ARM-forced simulations followed observations more closely than the default CLM4. There is still a warm bias in the minimum temperature range however. The diurnal temperature cycle in the summer shows a substantial warm bias in the maximum temperature calculations as well as a small warm bias in the calculation of minimum temperature. Both of these issues result in a large overestimation in the daily $T_{\rm skin}$ during the summer months. As noted earlier, LE is grossly erroneous during this period, and that may have at least some contribution to the daytime overestimation of T_{skin} . A comparison with the diurnal cycles of T_{air} and soil temperatures suggests that the model properly simulates the lag between the maximum and the minimum temperatures at the surface and between the maximum and the minimum temperatures in the soil. However, it is noted that the maximum July soil temperature in layer 3 is greater than T_{air} , suggesting that the warm bias in maximum T_{skin} leads to a warm bias in simulated soil temperatures.

The most urgent improvements that CLM4 needs may be of the surface energy fluxes, namely, the SH and the LE fluxes. It is likely that improvements to these fluxes will produce some improvement in T_{skin} calculations. In addition, the substantial increase in soil temperature error from layer 3 to layer 5 suggests that heat is not properly transported between those layers. Heat flux simulations might also benefit from improvements made to soil moisture, namely, in layer 5, which is often too moist.

Additional improvements in the calculation of snow cover and albedo at this site may be beneficial.

4. Conclusions

a. Future Work

Several uncertainties and key results of this study need to be addressed in future studies. The first uncertainty pertains to the land cover, the PFT structure, and the soil characteristics of this site. Although most of the ARM site is positioned over open pasture, the CO₂ flux site is directly over winter wheat. Furthermore, because soil color and composition were based on a coarse global data set, it is likely that they do not properly represent the soil type at the ARM site. This results in both the inaccurate drainage of moisture in the soil and the potentially inaccurate simulations of albedo. Determining a proper PFT mosaic and soil type for the site through sensitivity studies could allow for a more robust comparison between model simulations and observations. Furthermore, additional work needs to be performed to understand why the peaks in simulated soil moisture due to precipitation are underestimated compared with observations. Using a more accurate soil type and land cover may even eliminate this problem, or it could provide insight into other causes of the problem.

In addition, because the model is a one-dimensional simulation that does not incorporate horizontal and vertical wind components, SH fluxes are not properly simulated. Currently, the model sets the u (west to east component) to equal v (south to north component) based on a single wind speed. Being able to break that wind speed up into directional (i.e., u, v, and w) components would allow for a better simulation of SH fluxes. Future work using the ARM data could incorporate a directional wind component into CLM, which would lessen the uncertainty pertaining to the simulation of SH fluxes.

Furthermore, gaining insight into the cause of warm biases present in winter minimum and summer maximum temperatures is necessary. A potential cause is the inaccurate distribution of energy in the SEB, resulting in an overestimate of the amount of energy emitted as longwave radiation at certain times of the day. Understanding where the balance is inaccurately simulated and what can be modified would benefit the model. This would require sensitivity studies (similar to Zeng et al. 2012) to determine how energy needs to be properly distributed. In addition, the modifications would need to be shown to benefit model simulations globally rather than just at the ARM site.

Finally, further use of ARM data in evaluating CLM may be performed. The data set used here was only for one site and only for 1 year. Although this use of data is sufficient for validation studies, it does not take full advantage of the extensive amount of ARM observations available. The ARM project consists of six permanent ARM facilities and numerous mobile sites globally, with over one decade of observations at most sites (Xie et al. 2010). Furthermore, many of these sites take measurements at different locations in the site and for different purposes (such as the Climate Modeling Best Estimate data). Because ARM data proved to be useful in this evaluation, expanding its use to other locations on the same site, on other sites globally, and for other years would prove beneficial in further evaluations of CLM. Furthermore, using data taken from

different locations on the same site could increase the confidence in the results from this study.

b. Final Remarks

The purpose of this study was to perform an analysis on the offline CLM4, which is a critical component of the CESM climate model. When forced with ARM observations, many of the uncertainties related to the forcing data are reduced, allowing for a better understanding of how the model performs. The results demonstrated some of the issues currently affecting the CLM4 as well as the timescale in which they are most prevalent.

First, despite uncertainties, the ARM data introduced in this thesis performed exceptionally in its use as both forcing and evaluation data for the model. This was seen by both a similarity to offline CLM4 in terms of the diurnal and seasonal cycle of many of the surface energy variables and a general improvement in calculated errors for most of those when forced with the ARM-forcing data. The main variables that did not improve with the ARM-forcing data are those with known substantial issues in the model, specifically sensible and LE fluxes (Lawrence et al. 2011, 2012; Zeng et al. 2012).

Second, a diurnal and seasonal analysis of T_{skin} revealed a substantial warm bias in both winter minimum T_{skin} and summer maximum T_{skin} , which then ultimately leads to a warm bias in both of those seasons. This can be seen in both the seasonal scatterplots of offline CLM4 T_{skin} versus ARM T_{skin} (Figure 13) and the diurnal analysis of offline CLM4 versus ARM T_{skin} (Figure 17). The warm bias in winter minimum T_{skin} ultimately leads to a warm bias in many of the daily averaged calculations, although the bias is small. The substantial warm bias in summer maximum T_{skin} also leads to a daily warm bias in the averaged calculation, and thus, much of the summer warm bias is due to the warmer maximum temperatures. This warm bias is further explained when comparing the diurnal cycle of T_{skin} with T_{air} and soil temperature. Soil from layer 3 became warmer than maximum T_{air} in July, indicating that the warm bias in T_{skin} could potentially be due to absorption of heat by the soil. Because the summer warm bias is substantial, it contributes substantially to the calculated annual error for daily T_{skin} .

Third, this study showed that although energy and moisture are modeled to some accuracy below the surface, energy flux errors at the surface do indeed lead to energy flux errors in the soil below the surface. Soil temperature was found to have high precision and low bias in the upper layer, but higher (cold) bias in the lower layer. Soil moisture was found to be too dry in the upper layer and too moist in the lower layer, indicating that moisture fluxes in the model are too high (at least at this site). This can lead to errors in LE flux at the surface as well as in the thermal conductivity of the soil, which can then have profound impacts on T_{skin} .

This study identifies problems and suggests improvements to CLM4, which can lead to improved climate simulations using the CESM. Although many studies have used ARM data to validate atmospheric climate models, few studies have used ARM for the purpose of land modeling, and so this study shows that ARM data can be valuable for evaluating offline land models. Furthermore, this study revealed a warm bias at two different points of the diurnal cycle, demonstrating that accuracy on scales as small as even hourly can greatly affect the accuracy of the model daily, seasonally, and annually. There is great promise for the use of ARM data in further validating and improving land models.

REFERENCES

- Ackerman, T. P., and G. M. Stokes, 2003: The Atmospheric Radiation Measurement Program—to predict reliably what increased greenhouse gases will do to global climate, we have to understand the crucial role of clouds. *Phys. Today*, **56**, 38–46.
- Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, 90, 1283–1296.
- Anthes, R. A., 1984: Enhancement of convective precipitation by mesoscale variations in vegetative covering in semiarid regions. *J. Climate Appl. Meteor.*, **23**, 541–554.
- Baklanov, A. A., B. Grisogono, R. Bornstein, L. Mahrt, S. S. Zilitinkevich, P. Taylor, S.E. Larsen, M. W. Rotach, and H. Fernando, 2011: The nature, theory, and modeling of atmospheric planetary boundary layers. *Bull. Amer. Meteor. Soc.*, 92, 123–128.
- Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, and R. Evans, 2001: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.*, 82, 2415–2434.
- Bernstein, L., P. Bosch, O. Canziani, Z. Chen, R. Christ, O. Davidson, W. Hare, S. Huq, D. Karoly, and V. Kattsov, 2007: *Climate Change 2007: Synthesis Report*. Contribution of working groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change, Geneva. http://www.ipcc.ch/ipccreports/ar4-syr.htm.
- Bonan, G. B., 1996: Land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and users guide. NCAR Technical Note NCAR/TN-417 STR.
- Bonan, G. B., 2008: Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, **320**, 1444–1449.
- Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson, 2002: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochem. Cycles*, **16**, 1021.

- Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson, and Z. L. Yang, 2002: The land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model. J. Clim., 15, 3123– 3149.
- Brown, P. T., E. C. Cordero, and S. A. Mauget, 2012: Reproduction of twentieth century intradecadal to multidecadal surface temperature variability in radiatively forced coupled climate models. *J. Geophys. Res.*, **117**, D11116.
- Dai, Y., and Z. Qingcun, 1997: A land surface model (IAP94) for climate studies part I: Formulation and validation in off-line experiments. *Adv. Atmos. Sci.*, **14**, 433–460.
- Dai, Y., X. Zeng, R. E. Dickinson, I. Baker, G. B. Bonan, M. G. Bosilovich, A. S. Denning, P. A. Dirmeyer, P. R. Houser, and G. Niu, 2003: The common land model. *Bull. Amer. Meteor. Soc.*, 84, 1013–1023.
- Dickinson, R. E., 1995: Land processes in climate models. *Remote Sens. Environ.*, **51**, 27–38.
- Dickinson, R. E., 2000: How coupling of the atmosphere to ocean and land helps determine the timescales of interannual variability of climate. *J. Geophys. Res.*, **105**, 20115–20119.
- Dickinson, R., A. Henderson-Sellers, and P. Kennedy, 1993: Biosphere-Atmosphere Transfer Scheme (BATS) Version 1e as coupled to the NCAR community model. NCAR Technical Note NCAR/TN-387 STR, 72.
- Dirmeyer, P. A., 2000: Using a global soil wetness dataset to improve seasonal climate simulation. *J. Clim.*, **13**, 2900–2922.
- Dirmeyer, P. A., B. A. Cash, J. L. Kinter III, C. Stan, T. Jung, L. Marx, P. Towers, N. Wedi, J. M. Adams, and E. L. Altshuler, 2012: Evidence for enhanced landatmosphere feedback in a warming climate. *J. Hydrometeor.*, 13, 981–995.
- Fischer, M., 2005: Carbon Dioxide Flux Measurement Systems Handbook. Atmospheric Radiation Measurement Technical Report. ARM/TR048.
- Fischer, M. L., D. P. Billesbach, J. A. Berry, W. J. Riley, and M. S. Torn, 2007: Spatiotemporal variations in growing season exchanges of CO₂, H₂O, and sensible heat in agricultural fields of the Southern Great Plains. *Earth Interact.*, **11**, 1–21.
- Foken, T., 2008: The energy balance closure problem: An overview. *Ecol. Appl.*, **18**, 1351–1367.

- Frank, J. M., W. J. Massman, and B. E. Ewers, 2013: Underestimates of sensible heat flux due to vertical velocity measurement errors in non-orthogonal sonic anemometers. *Agric. For. Meteorol.*, **171**, 72–81.
- Gash, J. H. C., C. A. Nobre, J. M. Roberts, and R. L. Victoria, 1996: An overview of the Amazonian Climate Observation Study (ABRACOS). *Amazonian Deforestation and Climate*, J. H. C. Gash and et al., Eds., 1st ed. John Wiley and Sons, 1–14.
- Hall, D. K., and G. A. Riggs, 2007: Accuracy assessment of the MODIS snow products. *Hydrol. Process.*, 21, 1534–1547.
- Jin, M., 2004: Analysis of land skin temperature using AVHRR observations. *Bull. Amer. Meteor. Soc.*, **85**, 587–600.
- Jin, M., and R. E. Dickinson, 2010: Land surface skin temperature climatology: Benefitting from the strengths of satellite observations. *Environ. Res. Lett.*, 5, 044004.
- Jin, M., R. Dickinson, and A. Vogelmann, 1997: A comparison of CCM2-BATS skin temperature and surface-air temperature with satellite and surface observations. J. *Climate*, 10, 1505–1524.
- Jin, M., and S. Liang, 2006: An improved land surface emissivity parameter for land surface models using global remote sensing observations. J. Climate, 19, 2867– 2881.
- Jin, M. S., and T. J. Mullens, 2012: Land-biosphere-atmosphere interactions over the Tibetan plateau from MODIS observations. *Environ. Res. Lett.*, **7**, 014003.
- Jin, M., H. Bartholomew, and T. Mullens, 2013: Relations among soil moisture, soil temperature and surface temperatures from ARM observations and offline CLM4 simulations. *Environ. Res. Lett.* (submitted for publication).
- Lawrence, D. M., K. W. Oleson, M. G. Flanner, C. G. Fletcher, P. J. Lawrence, S. Levis, S. C. Swenson, and G. B. Bonan, 2012: The CCSM4 land simulation, 1850–2005: Assessment of surface climate and new capabilities. *J. Climate*, 25, 2240–2260.
- Lawrence, D. M., K. W. Oleson, M. G. Flanner, P. E. Thornton, S. C. Swenson, P. J.
 Lawrence, X. Zeng, Z. L. Yang, S. Levis, and K. Sakaguchi, 2011:
 Parameterization improvements and functional and structural advances in version
 4 of the Community Land Model. *J. Adv. Model. Earth Syst.*, **3**, M03001.

- Lawrence, P. J., and T. N. Chase, 2007: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). *J. Geophys. Res.*, **112**, G01023.
- Leuning, R., E. van Gorsel, W. J. Massman, and P. R. Isaac, 2012: Reflections on the surface energy imbalance problem. *Agric. For. Meteor.*, **156**, 65–74.
- Manabe, S., 1969: Climate and the ocean circulation 1. Mon. Wea. Rev., 97, 739–774.
- Neale, R. B., C. Chen, A. Gettelman, P. Lauritzen, S. Park, D. Williamson, A. Conley, R. Garcia, D. Kinnison, and J. Lamarque, 2010: Description of the NCAR Community Atmosphere Model (CAM 5.0). Technical Note NCAR/TN-486 STR.
- Nicholson, S. E., 1988: Land surface atmosphere interaction. *Prog. Phys. Geogr.*, **12**, 36–65.
- Niu, G. Y., and Z. L. Yang, 2007: An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. J. Geophys. Res., 112, D21101.
- Oleson, K. W., and G. B. Bonan, 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. *J. Hydrometeor.*, **1**, 431–446.
- Oleson, K. W., D. M. Lawrence, B. Gordon, M. G. Flanner, E. Kluzek, J. Peter, S. Levis, S. C. Swenson, E. Thornton, and J. Feddema, 2010: Technical description of version 4.0 of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-478 STR.
- Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson, 2006: Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations. J. *Hydrometeor.*, 7, 953–975.
- Qian, Y., C. Long, H. Wang, J. Comstock, S. McFarlane, and S. Xie, 2012: Evaluation of cloud fraction and its radiative effect simulated by IPCC AR4 global models against ARM surface observations. *Atmos. Chem. Phys.*, **12**, 1785–1810.
- Roesch, A., 2006: Evaluation of surface albedo and snow cover in AR4 coupled climate models. *J. Geophys. Res.*, **111**, D15111.
- Schlosser, C. A., 1996: Scales of temporal and spatial variability of midlatitude soil moisture. *J. Geophys. Res.*, **101**, 7163–7174.

- Sellers, P., R. Dickinson, D. Randall, A. Betts, F. Hall, J. Berry, G. Collatz, A. Denning, H. Mooney, and C. Nobre, 1997: Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, **275**, 502–509.
- Sellers, P., Y. Mintz, Y. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.*, **43**, 505–531.
- Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed. *Bull. Amer. Meteor. Soc.*, **75**, 1201–1221.
- Swenson, S., and D. Lawrence, 2012: A new fractional snow-covered parameterization for the Community Land Model and its effect on the surface energy balance. J. Geophys. Res., Atmos., 117, D21107.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res., 106, 7183–7192.
- Wan, Z., 2008: New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sens. Environ.*, **112**, 59–74.
- Wang, Z., and X. Zeng, 2010: Evaluation of snow albedo in land models for weather and climate studies. J. App. Meteor. Climatol., 49, 363–380.
- Williams, M., A. Richardson, M. Reichstein, P. Stoy, P. Peylin, H. Verbeeck, N. Carvalhais, M. Jung, D. Hollinger, and J. Kattge, 2009: Improving land surface models with FLUXNET data. *Biogeosciences*, 6, 1341–1359.
- Willmott, C. J., and K. Matsuura, cited 2000: Terrestrial air temperature and precipitation: Monthly and annual climatologies. <u>http://climate.geog.udel.edu/~climate</u>.
- Wiscombe, W. J., and S. G. Warren, 1980: A model for the spectral albedo of snow. I: Pure snow. J. Atmos. Sci., **37**, 2712–2733.
- Yang, F., K. Mitchell, Y. T. Hou, Y. Dai, X. Zeng, Z. Wang, and X. Z. Liang, 2008: Dependence of land surface albedo on solar zenith angle: Observations and model parameterization. J. App. Meteor. Climatol., 47, 2963–2982.
- Yang, Z. L., R. Dickinson, A. Henderson-Sellers, and A. Pitman, 1995: Preliminary study of spin up processes in land surface models with the first stage data of Project for Intercomparison of Land Surface Parameterization Schemes Phase 1 (a). J. Geophys. Res., 100, 16553–16578.

- Xie, S., R. B. McCoy, S. A. Klein, R. T. Cederwall, W. J. Wiscombe, E. E. Clothiaux, K. L. Gaustad, J-C. Golaz, S. H. Shamblin, and M. P. Jensen, 2010: ARM climate modeling best estimate data: A new data product for climate studies. *Bull. Amer. Meteor. Soc.*, 91, 13–20.
- Zeng, X., M. Shaikh, Y. Dai, R. E. Dickinson, and R. Myneni, 2002: Coupling of the common land model to the NCAR community climate model. J. Clim., 15, 1832– 1854.
- Zeng, X., Z. Wang, and A. Wang, 2012: Surface skin temperature and the interplay between sensible and ground heat fluxes over arid regions. *J. Hydrometeor.*, **13**, 1359–1370.

APPEDNIX A: Acronyms

ARM	Atmospheric Radiation Measurement
AVHRR	Advanced Very High Resolution Radiometer
CCM3	Version 3 of the Community Climate Model
CESM	Community Earth System Model
CLM	Community Land Model
G	Heat Flux into the Ground (or Ground Flux)
GCM	General Circulation Model
IPCC	Inter-Governmental Panel on Climate Change
LE	Latent Heat Flux
LSM1	Version 2 of the Land Surface Model
LSM2	Version 2 of the Land Surface Model
MODIS	Moderate Resolution Imaging Spectroradiometer
NCAR	National Center for Atmospheric Research
NCEP	National Center for Environmental Prediction
R	Pearson Correlation Coefficient

RMSE	Root-Mean Squared Error
SEB	Surface Energy Budget
SGP	Southern Great Plains
SH	Sensible Heat Flux
T _{air}	2-m Air Temperature
T _{skin}	Land Surface Skin Temperature
APPENDIX B: Improvements to Vegetation Canopy Emissivity

Adapted from T.J. Mullens et. al, 2013: Improving the Vegetation Canopy Parameterization in the Community Land Model (CLM4). Submitted to *Environmental Research Letters*.

Abstract

Vegetation Canopy Emissivity (ε_v) , an important parameter in calculating upward longwave radiation from the land surface, has been poorly represented in land surface models. For example, the Community Land Model (CLM4) calculates ε_v through a simple, empirical algorithm that leads to unrealistically low emissivity values for intermediate (0.5 < ELAI + ESAI < 2.5) and sparse canopy densities (ELAI + ESAI < 0.5). Such low ε_v causes CLM4 to underestimate upward longwave radiation and consequently overestimate land skin temperature (T_{skin}). This letter suggests a new parameterization that accounts for differences in emissivity by vegetation Plant Functional Type (PFT) to reduce the magnitude of the cavity effect in ε_v . Evaluation using the Atmospheric Radiation Measurement (ARM) ground observations show that the new parameterization improves CLM4 modeled daily T_{skin} by 0.17 °C on average. Furthermore, evaluation using satellite remote-sensing data shows that this new ε_V parameterization improves CLM4 monthly T_{skin} simulation up-to 1 °C for areas of intermediate canopy density. No evident improvements on T_{skin} simulations are found over sparse canopy or dense vegetation areas (ELAI+ESAI > 2.5) suggesting that this improvement is beneficial only to areas of intermediate canopy density.

B1. Introduction

The surface emissivity (ε) , which is the ratio of the emitted radiance from a surface over that of a blackbody at the same temperature, is an important parameter in land surface models for calculating upward longwave radiation. This variable, however, is hard to measure from satellite remote-sensing, partly because the sensor measures spectral emissivity which has to be converted into a broadband emissivity in order to be used in a land surface model (Jin and Liang 2006, here after JL06) and partly because satellites view the heterogeneous surface as one pixel while a land model may treat ground and vegetation separately. Originally, many land models assumed ε_v to be equal to 1 (Dickinson et al. 1986; Sellers et al. 1986), and some models currently still set this parameter close to or near 1 (this is referred to as the constant- ε approximation). Nevertheless, other land models now use a simple, first order approximation to calculate this parameter. For example, the Community Land Model version 4 (CLM4; Oleson et al. 2010, hereafter referred to as OEL10) calculates ground emissivity (ε_g) and vegetation emissivity (ε_v) independently. Specifically, ε_g in CLM4 is a function of fractional snow cover, bare soil emissivity, and snow emissivity, while ε_v is a function primarily of canopy structure, which is calculated as a sum of exposed one-sided leaf area index (ELAI) and exposed one-sided stem area index (ESAI):

$$\epsilon_{g} = (1 - f_{sno})\epsilon_{soil} - f_{sno}\epsilon_{sno} \text{ Eq. (B1) (OEL10, 4.24)},$$

 $\epsilon_{v} = 1 - e^{-(\text{ELAI} + \text{ESAI})/\mu} \text{ Eq. (B2) (OEL10, 4.25)},$

where $\varepsilon_{soil} = 0.96$ is the emissivity for bare soil, f_{sno} is the fraction of the ground covered by snow, $\varepsilon_{sno} = 0.97$ is the emissivity of snow, and $\mu = 1$ is the average inverse optical depth for longwave radiation (OEL10). ELAI is the ratio of total leaf area over the ground area covered by the plant, and ESAI is the ratio of the total area of all stems of a plant to the area of ground covered by the plant. This letter specifically focuses on improvements made to Eq. (B2). For a discussion on the accuracy and potential improvements to Eq. (B1), see JL06.

Because neither land nor vegetation are blackbodies, the emissivity parameters calculated above are then used to calculate total longwave radiation (L \uparrow), vegetation longwave radiation, (L_{vg} \uparrow) and skin temperature using the following equations:

$$L\uparrow = \delta_{\text{veg}}L_{\text{vg}}\uparrow + (1 - \delta_{\text{veg}})(1 - \varepsilon_g)L_{\text{atm}}\downarrow + (1 - \delta_{\text{veg}})\varepsilon_g \sigma(T_g^{\text{n}})^4 + 4\varepsilon_g\sigma(T_g^{\text{n}})^3(T_g^{\text{n}+1} - T_g^{\text{n}})$$

Eq. (B3) (OEL10 4.16),

where $L\uparrow$ is the total upward longwave radiation emitted from the land surface, $L_{atm}\downarrow$ is the downward longwave radiation emitted from the atmosphere, T_g is the temperature of the ground at time steps n and n+1, and δ_{veg} is 1 for vegetated surfaces and 0 for nonvegetated surfaces. Therefore, for vegetated surfaces ($\delta_{veg} = 1$), Eq. (B3) becomes:

$$L\uparrow = L_{vg}\uparrow + 4\epsilon_g \sigma(T_g^{n})^3(T_g^{n+1}-T_g^{n})$$
 Eq. (B4) (OEL10 4.18).

Because this letter focuses on vegetated areas, all values of L \uparrow are assumed to be calculated from Eq. (B4). L_{vg} \uparrow is upward longwave emission from vegetated surfaces, calculated using the equation:

$$L_{vg}\uparrow = (1 - \varepsilon_g)(1 - \varepsilon_v)(1 - \varepsilon_v)L_{atm}\downarrow + \varepsilon_v[1 + (1 - \varepsilon_g)(1 - \varepsilon_v)]\sigma(T_g^n)^3[T_v^n + 4(T_v^{n+1} - T_v^n)] + \varepsilon_g(1 - \varepsilon_v)\sigma(T_g^n)^4$$

Eq. (B5) (OEL10, 4.19),

where T_v is the temperature of the vegetation at time steps n and n+1. The first term in Eq. (B5) is radiation that is emitted by the atmosphere, transmitted down through the vegetation, reflected by the ground, and transmitted up through the vegetation. The second term is the radiation emitted directly from the canopy and the third term is radiation that is emitted upward from the ground and transmitted through the canopy and into the atmosphere (OEL10). Therefore, vegetation canopy emissivity affects the transmission of radiation from the atmosphere through the canopy, transmission and scattering of radiation from the ground, and emission of radiation from the canopy itself. The emission of radiation is determined by the scattering and emission of radiation from multiple walls and the floor, yielding an overall effect known as the "cavity effect" (Fuchs and Tanner 1996; Van De Griend and Owe 1993; Francois et al. 1997, Olioso et al. 2007 – hereafter referred to as OL07). The cavity effect is a result of the heterogeneity of the land surface (Prata et al. 1995). The total canopy emissivity is determined by the emissivity of the floor, roof, and the walls of the "cavity" (Prata et al. 1995). The total emissivity of the canopy increases when the ratio of the length of the canopy floor to its

height increases (OL07, Valor and Caselles 1996). A similar analogy is that of an urban canyon (Oleson et al. 2008). The urban land surface consists of floors and buildings with vertical walls and roofs. Each of these components has a different composition and thus a different emissivity. The total upward longwave radiation is determined by the sum of upward longwave radiation from each component, which is calculated by the temperature and emissivity of each surface (using the Stefan-Boltzmann law) and the reflected longwave radiation of each surface. Nevertheless, the total emissivity of the canyon increases with the ratio of the height (H) of the walls to the width of the canyon floor (W) because the total amount of emitted radiation from the walls increases with the height of the walls (Oleson et al. 2008). Therefore, erectophile canyons (those with high H-to-W ratios) have a higher canyon emissivity than more broad canyons (smaller H-to-W ratios). A further description of the urban canyon effect and its integration into CLM4 is found in Oleson et al. (2008).

 T_{skin} , the land surface skin temperature (Jin and Dickinson 2010) is then calculated from total L \uparrow based on the Stefan-Boltzmann law:

$$T_{skin} = \left(\frac{L\uparrow}{\sigma}\right)^{\frac{1}{4}}$$
 Eq. (B6) (OEL10, 4.15),

where σ is the Stefan-Boltzmann constant (5.67x10⁻⁸ Wm⁻²K⁻⁴). Therefore, accurate calculations of emissivity lead to improved calculations of L↑ and ultimately improved calculations of T_{skin}. The overestimation of T_{skin} in CLM4 is a long-standing problem for vegetated regions. This letter shows that the overestimation of T_{skin} can at least be partially attributed to the inaccurate calculation of ϵ_v In CLM4, the algorithm for ε_v produces unreasonably low values for vegetated surfaces in areas of lower canopy density (ELAI+SLAI < 2.5). For example, for cropland with an ELAI+ESAI between 0.5-1.5, Eq. (B2) produces a maximum ε_v of approximately 0.78 and a minimum of approximately 0.4, which are unrealistically low. These values then lead to underestimation of L \uparrow , which consequently leads to an overestimation of T_{skin} . This unrealistic ε_v reveals a shortcoming of the existing algorithm.

This letter proposes a new ε_v parameterization in Section 2. The sensitivity experiments are designed in Section 3. Section 4 shows the evaluations of the new ε_v parameterization by comparing CLM4 simulated T_{skin} with ARM and MODIS observations, followed by a brief discussion.

B2. Proposed Canopy Emissivity Algorithm

The emissivity of vegetation canopies depends on vegetation density, aerial extent and vegetation structure (e.g., height, leaf area index, etc.), which may differ across plant functional types (PFTs). In CLM4, vegetation cover is divided into 17 different PFTs, each having different optical properties, structure, and seasonality. For this study, a fixed emissivity value is applied to each PFT unless the literature suggests otherwise. The determined values for the PFT-based emissivity (ϵ_{PFT}) are listed in Table B1, and are based on a combination of literature review, and from MODIS averaged broadband emissivity values, calculated from Eq. 10 of JL06. Non-vegetated areas were assigned the default bare soil emissivity value of 0.96.

Table B1. Table of PFT Emissivity values determined by observations from Olioso e	et al
(2007), Rubio et al (1997), Valor and Caselles (1996), and Wittich (1997).	

PFT	ε _{max} (PFT)
Bare Soil (Non-Vegetated)	0.960
Needleleaf Evergreen Temperate Tree	0.982
Needleleaf Evergreen Boreal Tree	0.982
Needleleaf Deciduous Boreal Tree	0.985
Broadleaf Evergreen Tropical Tree	0.978
Broadleaf Evergreen Temperate Tree	0.981
Broadleaf Deciduous Tropical Tree	0.982
Broadleaf Deciduous Temperate Tree	0.970
Broadleaf Deciduous Boreal Tree	0.968
Broadleaf Evergreen Shrub	0.987
Broadleaf Deciduous Temperate Shrub	0.987
Broadleaf Deciduous Boreal Shrub	0.987
C3 Arctic Grass	0.978
C3 Non-Arctic Grass	0.978
C4 Grass	0.978
Corn	0.985
Spring Temperate Cereal	0.981
Winter Temperate Cereal	0.963
Soybean	0.977
Generic Crop	0.976
Irrigated Generic Crop	0.981

Because the cavity effect can have a significant effect on canopy emissivity (JL06; OL07), it is also important to include this effect in the calculation of canopy emissivity. Therefore, an algorithm based on the PFT emissivity values from Table B1 and the canopy density is proposed:

$$\varepsilon_{\rm v} = \varepsilon_{\rm PFT} - \delta \varepsilon * e^{-(\rm ELAI + ESAI)/\mu}$$
 Eq. (B7)

where $\mu = 1$ and $\delta \epsilon = 0.03$ (from OL07) is the maximum variation caused by the cavity effect, which only has a maximum magnitude of near 0.03, as suggested by OL07.

B3. Experimental Design

Two sets of sensitivity simulations are performed: First, a pair of single-point offline CLM4 simulations forced with ARM atmospheric observations and compared with ARM land observations for Lamont, OK. The second is a pair of global offline CLM4 simulations and the outputs are compared with the monthly NASA MODIS observed Tskin.

The two single-point offline CLM4 simulations are performed centered over the ARM Southern Great Plains (SGP) site in Lamont Oklahoma (36.6°N, 97.5°W) for the year 2004, with daily output. Both runs are forced by the ARM observed atmospheric data, including direct and diffuse solar radiation, downward longwave radiation, wind, precipitation and humidity, and evaluated using land observation data from the above experiment. A description of the ARM project can be found in Stokes and Schwartz. (1993), and a description of the data used can be found in Jin et al. (2013, submitted). Both runs are performed after a 50-year spin-up, with ARM-observed soil moisture as an initial condition, and the PFT's changed to a mosaic of 0.2 bare soil, 0.4 C3 grass and 0.4 C4 grass, determined from satellite photos and the ARM site description from Stokes and

Schwartz (1993). The first run ("ARM-forced Default") is performed with the default canopy emissivity equation (eq. B5) and the second run (ARM-forced PFT-emis) is performed with our new PFT-based emissivity equation (eq. B7).

To understand how the new ε_v affects the CLM at a global scale, the two global offline CLM4 simulations are conducted for the years 2001-2004, using a spatial resolution of 0.93° in latitude by 1.25° in longitude. The first simulation uses the default algorithm and the second uses the PFT-Based algorithm. Both global runs are forced using the default CLM4 forcing data of Qian et al. (2006). MODIS monthly T_{skin} measured on the Terra satellite during the years 2001-2004 is used to validate the global runs. The description and uncertainties of MODIS T_{skin} are discussed in Jin and Mullens (2012), and the accuracy is discussed in Wan (2008). The MODIS data has a resolution of 0.05° x 0.05° and is resized to match the resolution of 0.93° x 1.25° used in CLM4. Because MODIS produces a daytime (10:30 AM) and nighttime (10:30 PM) dataset for each month, these are averaged, producing a single monthly T_{skin} value at each point to compare with the monthly temperature values modeled by CLM4. The RMSE is then calculated at each grid point, similar to the single point case.

B4. Results

a. Evaluation of the new ε_v parameterization using ARM SGP observations

A time series of the calculated canopy emissivity values is given in Figure B1a for the default equation (Eq. B5) and for the new ε_v parameterization (Eq. B7). Apparently,

 ε_v is much more realistic after using the modified equation (Eq. B7) when compared to monthly MODIS broadband emissivity values.



Fig B1: (a) Satellite observed ELAI and ESAI for Lamont OK. (b) Daily default Canopy Emissivity values based on current emissivity parameterization (shown in blue) and proposed PFT-Based Canopy Emissivity values (shown in red) compared to MODIS monthly broadband values (green triangles). Values are for Lamont, Oklahoma for 2004.

The new ε_v improves the T_{skin} simulation in the offline CLM4 by 0.17°C (Figure B2). Forced with observed ARM forcing, the default ARM-forced CLM4 T_{skin} has an RMSE of 2.44°C, while with the new ε_v , the RMSE between PFT-emis CLM4 T_{skin} and ARM observed T_{skin} is reduced to 2.27°C (Figure B2a). Although both the default CLM run and the new ε_v run are warmer than the ARM observations by as much as 6°C, the new ε_v algorithm is less so (Figure B2c), which is a reduction of the warm bias. Late fall, winter, and early spring are the most sensitive to the change (Figure B2c), which is when canopy density is lowest; this suggests that lower canopy densities are indeed most sensitive to the new algorithm. Although the algorithm reduces the warm bias, it does not substantially affect either the day-to-day cycle of Tskin changes, or the day-to-day changes in difference between modeled and observed T_{skin} at the ARM SGP site; it merely reduces the magnitude of the departure from observations. In addition, the difference taken between the two daily outputs (Figure B2c) shows that the late fall to early spring periods are most sensitive to ε_v improvement, suggesting that canopies with lower densities are improved most.



Figure B2 a) Offline ARM-Forced CLM4 Tskin with emissivity based on current parameterization and Offline ARM-Forced CLM4 Tskin with proposed emissivity calculated by equation B5 versus ARM-observed Skin Temperatures for the Southern Great Plains ARM Facility in 2004. b) Difference between CLM4 model runs and ARM land observations for 2004. c) Calculated difference between Offline ARM-forced Default and Offline ARM-forced PFT-Emissivity CLM4 runs.

In addition to daily changes to T_{skin} caused by the implementation of Eq. (B7), the changes also exhibit a diurnal cycle (Figure B3). For example, in July, the implementation of the new parameterization leads to cooling of as much as $0.7 \,^{\circ}$ C in the morning as T_{skin} is increasing and warming of as much as $0.5 \,^{\circ}$ C in the late afternoon as T_{skin} begins to decrease (Figure B3b). A possible cause of this result is a decrease in the rate of warming and cooling due to the new parameter. The cooler temperatures in the morning indicate that the new emissivity slows down the warming process and the warmer temperatures in the late afternoon indicate that the new emissivity slows down the cooling process. It takes longer for the canopy to respond to changes in incoming radiation. Further studies need to be performed to truly understand the cause of this decrease in the morning in the morning in the morning in the afternoon.



Figure B3 a) Offline hourly ARM-Forced CLM4 Tskin with emissivity based on current parameterization and Offline ARM-Forced CLM4 Tskin with proposed emissivity calculated by equation B5 versus ARM-observed Skin Temperatures for the Southern Great Plains ARM Facility in July, 2004. c) Calculated difference between Offline ARMforced Default and Offline ARM-forced PFT-Emissivity CLM4 runs.

b. Evaluation of the new ε_v parameterization using MODIS data

The distribution of ELAI + ESAI in CLM4 (Figure B4a) suggests that many regions may be affected by the new ε_v parameterization since most of the land has low to intermediate values of ELAI + ESAI. The difference in RMSE between the new ε_v case and default case shows encouraging improvements in T_{skin} simulations (Figure B4b). Areas of blue indicate a decrease in error when the new ε_{v} algorithm is used while areas of red indicate that the error increased with use of the new algorithm. Specifically, areas of intermediate canopy density (with ELAI+ESAI in the range of 0.5 to 2.5 as shown in Figure B4a) show improvement as high at 1 K. These areas include the Southwestern edge of the Tibetan Plateau in Northern India, savanna areas between the Sahara desert and the Gulf of Guinea, while areas such as the Indian Peninsula, interior Southern Africa, and the Southern fringe of the Boreal forests in Northern Europe experience a marginal improvement of 0.2° C to 0.5° C, similar to the improvement seen over the ARM SGP site. Error increases occur over arid and semi-arid areas such as the Australian deserts, the southeastern edge of the Tibetan Plateau, parts of the Taklimakan desert, the Siberian tundra on the eastern Siberian peninsula, and the western United States Mountains. Areas of heavily dense canopy (ELAI+ESAI above 2.5) experience little noticeable change in RMSE between the two simulations.



Figure B4. (a) Map of monthly Canopy Density (ELAI+ESAI) used as CLM4 input, averaged over 2004. The ELAI and ESAI values are CLM4 surface conditions, based on MODIS observations. (b) Difference in RMSE between default monthly offline CLM4 run and CLM4 run with improved vegetation emissivity. Average MODIS monthly Tskin for 2001-2004 is used as observational data. MODIS data is average of Monthly daytime and nighttime data collected by the MODIS instrument on the Terra Satellite. MODIS data has a resolution of 0.05 x 0.05 degrees, which is reduced to the 0.93 x 1.25 resolution of the CLM4 grid.

B5. Discussion

The introduction of Eq. (B7) to the CLM proves beneficial for the calculation of T_{skin} over intermediate canopies. The ARM SGP site is a winter wheat site with seasonally variable canopy density and has T_{skin} improvements of 0.17°C. Areas such as the Southwestern edge of the Tibetan Plateau and the African Savanna have greater seasonality of canopy density due to monsoonal presence and the shifting in the Inter-Tropical Convergence Zone over Africa. These regions of intermediate canopy density currently have unreasonably low ε_v values based on the default algorithm, and show significant improvement by up to 1°C in T_{skin} simulation. Areas with sparse canopy have the biggest increase in RMSE, due to overestimation of emissivity in these areas. This problem is noted in JL06, where bare soil emissivity values below 0.9 were suggested to be reasonable. In addition, many of these areas have recorded cold biases (Zeng et al. 2012); therefore the implementation of a higher vegetation emissivity in these areas can add to that cool bias. However, many of these areas are more subject to bare soil emissivity because they have little vegetation, so improvements to the bare soil algorithm might solve problems in these areas. This research suggests that further work is needed in developing an algorithm for sparse canopies, with a focus on bare ground emissivity, which can be integrated into Eq. (B7) to produce globally accurate emissivity values in CLM4.

77

Acknowledgements

This work is funded by THE NSF atmosphere program through grant (AWARD NUMBER ATM 0855480). The authors would also like to acknowledge the National Center for Atmospheric Research's Advanced Studies Program/Graduate Visitor Program, which allowed for substantial amounts of this work to be carried out at NCAR.

References

- Dickinson, R.E., A. Henderson-Sellers, P. J. Kennedy, and M. F. Vilson, 1986: Biosphere-Atmosphere Transfer Scheme (BATS) for the NCAR community climate model.NCAR Tech. Note TN-275+STR, NCAR, Boulder, CO, 48 pp.
- Francois, C., C. Ottlé, and L. Prévot, 1997: Analytical parameterization of canopy directional emissivity and directional radiance in the thermal infrared. Application on the retrieval of soil and foliage temperatures using two directional measurements. *Int.J.Remote Sens.*, 18, 2587-2621.
- Fuchs, M., C. Tanner, 1966: Infrared thermometry of vegetation. Agron. J., 58, 597-601.
- Jin, M., H. Bartholomew, and T. Mullens, 2013: Relations among soil moisture, soil temperature and surface temperatures from ARM observations and offline CLM4 simulations. *Env. Res. Lett.* Submitted.
- Jin, M., R. E. Dickinson, and A. M. Vogelmann, 1997: A comparison of CCM2/BATS skin temperature and surface-air temperature with satellite and surface observations. J. Climate, 10, 1505-1524.
- Jin, M., and S. Liang, 2006: An Improved Land Surface Emissivity Parameter for Land Surface Models Using Global Remote Sensing Observations. J. Climate, 19, 2867-2881.
- Jin, M., and T. J. Mullens, 2012: Land-Biosphere-Atmosphere Interactions Over Tibetan Plateaufrom MODIS Observations. *Env. Res. Lett.*, **7**, 9 pp.
- Oleson, K. W., D. M. Lawrence, B. Gordon, M. G. Flanner, E. Kluzek, J. Peter, S. Levis, S. C. Swenson, E. Thornton, and J. Feddema, 2010: Technical description of version 4.0 of the Community Land Model. NCAR Tech. Note. NCAR/TN-478+STR, 257 pp.
- Oleson, K. W., G. Bonan, J. Feddema, M. Vertenstein, and C. Grimmond, 2008: An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. *Journal of Applied Meteorology and Climatology*, 47, 1038-1060.
- Olioso, A., G. Soria, J. Sobrino, and B. Duchemin, 2007: Evidence of Low Land Surface Thermal Infrared Emissivity in the Presence of Dry Vegetation. *IEEE Geosci. and Rem. Sens. Lett.*, **4**, 112-116.

- Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson, 2006: Simulation of global land surface conditions from 1948 to 2004: Part I: Forcing data and evaluations. *J.Hydrometeor.*, 7, 953-975.
- Rubio, E., V. Caselles, and C. Badenas, 1997: Emissivity measurements of several soils and vegetation types in the 8-14 μm wave band: Analysis of two field methods. *Remote Sens. Env.*, **59**, 490-521.
- Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. J. Atmos. Sci., 43, 505-531.
- Stokes, G. M., and S. E. Schwartz, 1994: The atmospheric radiation measurement (ARM) program: Programmatic background and design of the cloud and radiation test bed. *Bull. Amer. Meteor. Soc.*, **75**, 1201-1221.
- Valor, E., and V. Caselles, 1996: Mapping land surface emissivity from NDVI: Application to European, African, and South American Areas. *Remote Sens. Env.*, 57, 167-184.
- Van de Griend, A., M. Owe, 1993: On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. *Int. J. Remote Sens.*, 14, 1119-1131.
- Wan, Z. 2008: New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. *Remote Sensing of Environment*, **112**, 59-74.
- Wittich, K.-P., 1997: Some simple relationships between land-surface emissivity, greenness and the plant cover fraction for use in satellite remote sensing. *Int J. Biometeorol.*, **41**, 58-64.
- Xie, S., R. B. McCoy, S. A. Klein, R. T. Cederwall, W. J. Wiscombe, E. E. Clothiaux, K. L. Gaustad, J-C. Golaz, S. H. Shamblin, and M. P. Jensen, 2010: ARM climate modeling best estimate data: A new data product for climate studies. *Bull. Amer. Meteor. Soc.*, **91**, 13-20.