Generation of an empirical soil moisture initialization and its potential impact on subseasonal forecasting skill of continental precipitation and air temperature

Marie Boisserie
GENERATION OF AN EMPIRICAL SOIL MOISTURE INITIALIZATION AND ITS POTENTIAL IMPACT ON SUBSEASONAL FORECASTING SKILL OF CONTINENTAL PRECIPITATION AND AIR TEMPERATURE

By

MARIE BOISSERIE

A Dissertation submitted to the Department of Meteorology in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Degree Awarded:
Spring Semester, 2010
The members of the committee approve the dissertation of Marie Boisserie defended on January 26, 2010.

________________________
James O’Brien
Professor Directing Dissertation

________________________
Eric Chassignet
University Representative

________________________
Steven Cocke
Committee Member

________________________
Paul Ruscher
Committee Member

________________________
Xiaolei Zou
Committee Member

Approved:

Robert Ellingson, Chair, Department of Meteorology

Joseph Travis, Dean, College of Arts and Sciences

The Graduate School has verified and approved the above-named committee members.
to Flavien, my family and my friends ...
ACKNOWLEDGEMENTS

The writing of a dissertation can be a lonely and isolating experience, yet it is obviously not possible without the personal and practical support of numerous people. Thus, in this section, I would like to express my gratitude to all people who helped me in one way or another.

I first would like to express my deep and sincere gratitude to my supervisor, Professor Emeritus James O’Brien, Ph.D., former head of the Center for Ocean-Atmospheric Prediction Studies (COAPS), Florida State University, for bringing me to COAPS as an intern and keeping me as a PhD student. His wide knowledge and experience have been of great value for me. It was a great privilege and honor for me to be his student.

I am deeply grateful to Dr. Steve Cocke for his personal guidance, and his constructive scientific comments throughout my dissertation. Besides being of great help in the evolution of my research, his encouragement, friendship and support have cheered me up in difficult moments. Without his encouragement and constant guidance I could not have finished this dissertation.

I wish to express my warm and sincere thanks to all my committee members, Professor Eric Chassignet, Dr. Steve Cocke, Professor Xiaolei Zou and Professor Paul Ruscher, who provided me with constructive criticism, good questions and excellent advice during the preparation of this dissertation.

I owe a special gratitude to my outside committee member, Professor Eric Chassignet for sharing his global view of what is the work of a scientist as opposed to that of an engineer.
He also advised me on how to write a scientific paper and how to give a talk in a constructive and systematic manner.

I warmly thank Dr. Vasu Misra, who is responsible for involving me in an international research project, the Global Land-Atmosphere Coupled Experiment (GLACE-2). Participating in this international research group gave me the opportunity to work with an international team and gave me the honorable responsibility of representing COAPS/FSU in the field of land-atmosphere coupling studies.

I would like to thank my climate team at COAPS, Tim LaRow, Dong-Wook Shin, Lim Young-Kwon and the last but not least Lydia Stevanova. They all gave me scientific and technical advice in Grads and/or Fortran programming. I will always remember our weekly meeting at Momo’s pizza. Special thanks to Shin who showed to me in the first place how to compile and run the coupled land-atmosphere FSU/COAPS model that I have used to produce my numerical experiments. Special thanks also to Lydia for her very pertinent questions, advice and for being a nice, interesting and amusing person.

My last gratitude is to all the people from COAPS for making this center a pleasant and friendly work environment.
# TABLE OF CONTENTS

List of Figures ................................................................. vii
List of Tables ................................................................. xi
Abstract ................................................................. xiii

1. INTRODUCTION ................................................................. 1

2. BACKGROUND ................................................................. 3

3. DEVELOPMENT OF A SOIL MOISTURE ANALYSIS ......................... 7
   3.1 Model ................................................................. 8
   3.2 PAR Technique ........................................................ 8
   3.3 Soil Moisture Datasets ............................................... 14
   3.4 Experimental Design ............................................... 17
   3.5 Results ................................................................. 18
   3.6 Summary and Discussion .......................................... 39

4. THE IMPACT OF A NEW SOIL MOISTURE ANALYSIS ON SHORT-TERM
   TO SUBSEASONAL FORECASTING SKILL .................................. 41
   4.1 GLACE-2 Overview ............................................... 41
   4.2 Experimental Design ............................................... 44
   4.3 Observational Datasets ........................................... 47
   4.4 Results ................................................................. 47
   4.5 Summary and Discussion .......................................... 74

5. CONCLUSION ................................................................. 76

APPENDICES .......................................................................... 80

A. SOIL MOISTURE MODEL .................................................. 80

REFERENCES ................................................................. 83

BIOGRAPHICAL SKETCH ....................................................... 90
LIST OF FIGURES

1.1 Theory of the land-atmosphere feedbacks. The negative sign indicates a decrease and the positive sign indicates an increase of the variable considered. 2

3.1 Schematic of the PAR technique. .................................................. 10

3.2 Average number of gauges in 2.5° grid box per month between January 1986 and December 1995. .............................................................. 13

3.3 In-situ soil moisture stations (crosses) over Illinois. The red grid represents the grid points of the FSU/COAPS model. ................................. 16

3.4 Two numerical simulations using the land-atmosphere FSU/COAPS model. .............................................................. 17

3.5 Global soil moisture average of the deepest layer (229 to 343 soil depth) during the last 8 years of the 50-year spinup time period. .................. 18

3.6 Average precipitation (mm/month) over the a) boreal summer months and b) boreal winter months from 1986 to 1995. ................................. 20

3.7 Statistical tools calculated against the daily totals of GDMF for the a) boreal summer (JJA) and b) boreal winter (DJF). All grid points of the global scale are taken into account. .................................................. 21

3.8 Same as Figure 3.7 but against the monthly GPCC dataset. ................ 22

3.9 Monthly soil moisture climatology of the in-situ observation (thick grey), CONTROL (dotted), PAR (dashed), R2 (semi-dotted) and GSWP2 (thin grey) in a) the top 10 cm layer and b) the deeper 10-200 cm layer at 6 locations in Illinois. .............................................................. 25

3.10 Vertical profile of soil moisture anomalies for the period 1986-1990. ....... 27


3.12 Monthly soil moisture anomaly correlation a) in the top 10 cm layer and b) in the deep layer 10-200 cm at 6 different locations in Illinois. .......... 28
3.13 Monthly soil moisture a) seasonal average and b) anomaly correlation against the in-situ observation with respect to 1986-1995 time period in the top 1.5 cm layer averaged over Illinois. For the plot of seasonal average, the thick grey line indicates the in-situ observations, the dotted line indicates the CONTROL simulation, the dashed line indicates the PAR simulation, the semi-dotted line indicates R2 and the thin black line indicates GSWP2.

3.14 Climatology (1st column), standard deviation (2nd column) and anomaly correlation of the average monthly a) precipitation, b) surface runoff and c) soil evaporation across Illinois (1986-1995).

3.15 Average soil moisture (mm/month) from 1986 to 1995 of a) the boreal summer months and b) the boreal winter months.

3.16 Average surface runoff (mm/month) from 1986 to 1995 of a) the boreal summer months and b) the boreal winter months.

3.17 Average surface evaporation (mm/month) from 1986 to 1995 of a) the boreal summer months and b) the boreal winter months.

3.18 Time anomaly correlation between each numerical simulation and observations for a) soil moisture b) surface runoff and c) surface evaporation across the years from 1986 and 1995.

3.19 Surface water budget as defined in equation 3.4 of a) the GSWP-2, b) the R2, c) the PAR and d) the CONTROL average from 1986 to 1995.

4.1 Land-atmosphere coupling from GLACE-1 by Koster et al., [2004]. The black boxes are the so-called "hotspots", regions where coupling strength is high.

4.2 Vertical profile of average soil moisture anomalies over Illinois for the observations, the PAR forecasts (real. for) and the control forecasts (ctl for). 7 layers are shown: top-10, 10-30, 30-50, 50-70, 70-90, 90-110 and 110-200 cm. The left hand side represents the anomaly correlation between each set of forecasts and the observations. The gray area indicates that the anomaly is not statistically significant at the 99% confidence interval.

4.3 2-m air temperature anomaly correlations against observations of the PAR forecasts (real. for, left column) and the control forecasts (ctl for, middle column) for 15 days, 1-month and 2-month forecasts. The right column displays the percentage of skill increase (or decrease) in terms of fraction of explained variance attributed to the realistic soil moisture initialization.
4.4 Time series of the average forecasted anomalies of T2m (top panel) and soil moisture vertical profile of PAR forecasts (real. for, middle panel) and the control forecasts (ctl for, bottom panel) over the box centered in grid cell 5 (Figure 4.6a). The time scale of the forecasts considered here is the average over the first month.

4.5 Same as Figure 4.3 but for precipitation forecasts.

4.6 Locations of the grid cells (black crosses) and areas (black boxes) overlaying the difference of the time anomaly correlations square between the PAR and control forecasts at 1-month time scale for a) air temperature and b) precipitation forecasts displayed in Figures 4.3 and 4.5 respectively.

4.7 Relationship between a change of the forecasted 2-m air temperatures and a change of the forecasted soil moisture (sm). Delta(sm) and delta(T2m) are defined in equation 4.3.

4.8 Relationship between a change of the forecasted precipitation (P) and a change of the forecasted soil moisture (sm). Delta(sm) and delta(P) are defined in equation 4.3.

4.9 2-m air temperature anomaly correlation square ($r^2$) against observations of the series-2 (left column) and the series-1 (middle column) for 1-15 days time scale. The right column displays the difference between the series-1 and series-2. The series-1 is equivalent to our PAR forecasts and the series-2 is equivalent to our control forecasts.

4.10 Same Figure as 4.9 but for 16-30 time scale.

4.11 Same Figure as 4.9 but for 31-45 time scale.

4.12 Same Figure as 4.9 but for 46-60 days time scale.

4.13 Precipitation anomaly correlation square ($r^2$) against observations of the series-2 (left column) and the series-1 (right column) for 1-15 days time scale. The right panels displays the difference between the series-1 and series-2. The series-1 is equivalent to our PAR forecasts and the series-2 is equivalent to our control forecasts.

4.14 Same Figure as 4.13 but for 16-30 days time scale.

4.15 Same Figure as 4.14 but for 31-45 days time scale.

4.16 Same Figure as 4.15 but for 46-60 days time scale.
4.17 Average precipitation (top panels), the 2-m air temperature (middle panels) and the total soil moisture anomalies (bottom panels) of the PAR forecasts (real for, left column), the control forecasts (ctl for, middle column) and the observations (right column) for the summer months (JJA) of the year 1988. The precipitation anomalies are in mm.d$^{-1}$, the 2-m air temperature are in K.d$^{-1}$ and the volumetric soil moisture m$^3$.m$^{-3}$.month$^{-1}$.

4.18 Same as Figure 4.17 but for the flood year of 1993.

4.19 Anomalies difference between the PAR forecasts and the control forecasts of precipitation (top), 2-m air temperature (middle) and soil moisture (bottom) for a) the 1988 drought (left column) and b) the 1993 flood (right column).

4.20 Evaporative sources (kg m$^{-2}$) for rain falling over the target region (black rectangle) during June-August of the year a) 1988 and b) 1993. The difference of evaporation sources between the year 1988 and 1993 is displayed in panel c.

A.1 CLM2 Soil Moisture Model.
LIST OF TABLES

4.1 List of participants in the GLACE-2 experiment. 44

4.2 Summary of the experiments: name of the series of forecasts, its description, number of forecasts, study period, time scales of the forecasts, and equivalent name used in the GLACE-2 experiment. 45

4.3 Spatial correlations squared ($r^2$) over the U.S. between each series of forecasts and the observations for each boreal summer month of the drought year of 1988. The difference of $r^2$ between the PAR forecasts and the control forecasts was also computed. The values of correlation or the difference of correlation are in bold when they are statistically significant at the 95% confidence interval. 69

4.4 Same as Table 4.3 but the flood year of 1993. 70
LIST OF ABBREVIATIONS

DOE  Department Of Energy
CLM2  Community Land Model Version 2
CLM3  Community Land Model Version 3
GDMF  Global Dataset of Meteorological Forcing for land surface
GLACE-2  Global Land Atmosphere Coupling Experiment 2nd Phase
GPCC  Global Precipitation Climatology Center
GSWP-2  Global Soil Wetness Project 2nd Phase
FSU/COAPS  Florida State University/Center for Ocean-Atmospheric Predictions Studies
ISLSCP-II  International Surface Climatology Project second-phase
LSM  Land Surface Model
NCAR  National Centers for Atmospheric Research
NCEP  National Centers for Environmental Predictions
OSU  Oregon State University
R2  NCEP/DOE reanalysis 2
PAR  Precipitation Assimilation Reanalysis
TRMM  Tropical Rainfall Measuring Mission
ABSTRACT

The goal of this dissertation research is to produce empirical soil moisture initial conditions (soil moisture analysis) and investigate its impact on the short-term (2 weeks) to subseasonal (2 months) forecasting skill of 2-m air temperature and precipitation. Because of soil moisture has a long memory and plays a role in controlling the surface water and energy budget, an accurate soil moisture analysis is today widely recognized as having the potential to increase summertime climate forecasting skill. However, because of a lack of global observations of soil moisture, there has been no scientific consensus on the importance of the contribution of a soil moisture initialization as close to the truth as possible to climate forecasting skill. In this study, the initial conditions are generated using a Precipitation Assimilation Reanalysis (PAR) technique to produce a soil moisture analysis. This technique consists mainly of nudging precipitation in the atmosphere component of a land-atmosphere model by adjusting the vertical air humidity profile based on the difference between the rate of the model-derived precipitation rate and the observed rate.

The unique aspects of the PAR technique are the following: 1) based on the PAR technique, the soil moisture analysis is generated using a coupled land-atmosphere forecast model; therefore, no bias between the initial conditions and the forecast model (spinup problem) is encountered; and 2) the PAR technique is physically consistent; the surface and radiative fluxes remains in conjunction with the soil moisture analysis. To our knowledge, there has been no attempt to use a physically consistent soil moisture land assimilation system into a land-atmosphere model in a coupled mode.
The effect of the PAR technique on the model soil moisture estimates is evaluated using the Global Soil Wetness Project Phase 2 (GSWP-2) multimodel analysis product (used as a proxy for global soil moisture observations) and actual in-situ observations from the state of Illinois. The results show that overall the PAR technique is effective; across most of the globe, the seasonal and anomaly variability of the model soil moisture estimates well reproduce the values of GSWP-2 in the top 1.5 m soil layer; by comparing to in-situ observations in Illinois, we find that the seasonal and anomaly soil moisture variability is also well represented deep into the soil. Therefore, in this study, we produce a new global soil moisture analysis dataset that can be used for many land surface studies (crop modeling, water resource management, soil erosion, etc.).

Then, the contribution of the resulting soil moisture analysis (used as initial conditions) on air temperature and precipitation forecasts are investigated. For this, we follow the experimental set up of a model intercomparison study over the time period 1986-1995, the Global Land-Atmosphere Coupling Experiment second phase (GLACE-2), in which the FSU/COAPS climate model has participated. The results of the summertime air temperature forecasts show a significant increase in skill across most of the U.S. at short-term to subseasonal time scales. No increase in summertime precipitation forecasting skill is found at short-term to subseasonal time scales between 1986 and 1995, except for the anomalous drought year of 1988.

We also analyze the forecasts of two extreme hydrological events, the 1988 U.S. drought and the 1993 U.S. flood. In general, the comparison of these two extreme hydrological event forecasts shows greater improvement for the summertime of 1988 than that of 1993, suggesting that soil moisture contributes more to the development of a drought than a flood. This result is consistent with Dirmeyer and Brubaker [1999] and Weaver et al. [2009]. By analyzing the evaporative sources of these two extreme events using the back-trajectory methodology of Dirmeyer and Brubaker [1999], we find similar results as this latter paper; the soil moisture-precipitation feedback mechanism seems to play a greater role during the drought year of 1988 than the flood year of 1993.
Finally, the accuracy of this soil moisture initialization depends upon the quality of the precipitation dataset that is assimilated. Because of the lack of observed precipitation at a high temporal resolution (3-hourly) for the study period (1986-1995), a reanalysis product is used for precipitation assimilation in this study. It is important to keep in mind that precipitation data in reanalysis sometimes differ significantly from observations since precipitation is often not assimilated into the reanalysis model. In order to investigate that aspect, a similar analysis to that we performed in this study could be done using the 3-hourly Tropical Rainfall Measuring Mission (TRMM) dataset available for the time period 1998-present. Then, since the TRMM dataset is a fully observational dataset, we expect the soil moisture initialization to be improved over that obtained in this study, which, in turn, may further increase the forecast skill.
CHAPTER 1

INTRODUCTION

The possible modifications of climate (i.e. climate change) on continents are a major concern for societies world-wide. An integral part of the earth’s climate system is the global hydrological cycle. The change and variability of the global hydrological cycle have great implications for agriculture, economy and human life. Water-related natural disasters, in particular hurricanes, floods, droughts, heatwaves, have been more devastating as far as deaths, suffering, and economical damages, than other natural disasters (i.e. earthquakes, volcanoes, etc.) [Kundzewicz et al., 1993]. For instance, over the past 30 years, the U.S has experienced 90 natural disasters accounting for $700 billion in total losses and more than 8,000 deaths. In Europe, the 2003 heatwave was one of the hottest summers on record, especially in France, and killed more than 38,000 people in total.

Although the soil moisture amount seems to be insignificant when compared to the total amount of water on a global scale, this hydrological variable is today widely recognized as crucial for climate predictions including extreme events. Soil moisture is defined as the moisture in the top several meters of soil that interacts with the atmosphere. It is a key land variable because it is slow-varying and can thus be predicted up to two months in advance (particularly at deep soil layers), and its role in controlling the exchange of water and heat transfer between the land surface and the atmosphere through evapotranspiration. During the summertime, when soil evaporation is strong, the water stored in the soil is released back into the atmosphere. This cools the soil and increases the relative humidity of the air. Changes in relative humidity of the air can affect many near surface variables, such as air temperature, surface pressure, atmospheric circulation, albedo and thus radiation budget. Through its effect on moisture convergence over the land surface, soil moisture also
has the potential to trigger the generation of precipitation. Then, a feedback mechanism between soil moisture and precipitation takes place (Figure 1.1). A significant soil moisture-precipitation feedback is thought to be responsible for the enhancement of the prolongation and/or intensity of the hydrological natural disasters (i.e., drought, flood, or heatwave) [Pal and Eltahir, 2001; Sud et al., 2003]. Therefore, improved initial state of soil moisture is expected to increase climate and extreme event forecasting skill.

Figure 1.1: Theory of the land-atmosphere feedbacks. The negative sign indicates a decrease and the positive sign indicates an increase of the variable considered.
CHAPTER 2

BACKGROUND

Many numerical studies have shown a significant sensitivity of near-surface climatological variables to soil moisture levels [Shukla and Mintz, 1982; Rind et al., 1982; Yeh et al., 1984; Sud and Fenessy, 1984; Koster et al., 2000; Hong and Kalnay, 2000, among others]. More recent numerical studies have shown that, during the boreal summer, while the continental precipitation variability is mostly influenced by Sea Surface Temperatures (SSTs) in the tropics, it is mostly influenced by soil moisture in the mid-latitudes [Kumar and Hoerling, 1995; Trenberth et al., 1988; Shukla, 1998; Koster et al., 2000]. Consequently, all the above sensitivity studies have demonstrated that accurate soil moisture initial conditions can potentially improve subseasonal forecasts of near surface variables. However, most of these studies use extreme values of soil moisture initial conditions (almost dry or saturated). The literature dealing with the use of soil moisture initial conditions that are as close to the truth as possible is not very extended. Progress in addressing this question has been hampered by the lack of reliable observational global soil moisture data sets to initialize global climate models. Indeed, the heterogeneity of the soil properties (e.g. porosity, permeability), the topography and the land cover types, make a global soil moisture measurement difficult. Today, this variable is sparsely measured in-situ and is not well estimated by satellite remote sensing. Despite their significant advances, the current remote sensing techniques for soil moisture still suffer from issues associated with the shallow depth of the retrieval (less than 2 cm), the absence of retrieval over dense vegetated and frozen areas, and significant uncertainties in the retrieval algorithm. International initiatives have provided long-term global soil moisture estimates by combining different off-line Land Surface Models (LSMs) driven by observation-based atmospheric forcing, such as the 2nd Global Soil Wetness Project [GSWP-2, Dirmeyer et al., 2002], and more recently the Global Land Data Assimilation
(GLDAS) system and the North American Land Data Assimilation (NL-DAS) system. Since those off-line LSMs use different land surface schemes, their soil moisture analysis product cannot be used for initialization but can be used as an evaluation dataset for land surface estimates.

In an ongoing model intercomparison project named the 2nd Global Land Atmosphere Coupling Experiment (GLACE-2), almost all the participants drive their LSM in a offline mode with the GSWP-2 observation-based atmospheric forcing data. The GSWP-2 is one of the state-of-the-art atmospheric forcing data sets for land surface analysis available in the GLACE-2 study period (1986-1995). The land surface state variables from the offline simulation are then used to initialize the GLACE-2 coupled land-atmosphere model. However, because the offline simulation and the land-atmosphere model have most likely a different climatology, the near surface atmospheric state of the forecasts may undergo an adjustment (or spinup). This spinup problem can decrease the short-term to subseasonal forecast skill of near surface variables. A land assimilation system integrated into a coupled land-atmosphere model is expected to reduce this spinup problem. We have recently joined the GLACE-2 team and but we use a different land assimilation system integrated into a land-atmosphere in a coupled mode (explained below).

Starting in 2002, reanalysis products provided by operational centers were among the first to use a land assimilation system integrated into a land-atmosphere model in a coupled mode. For instance, the National Center for Environmental Prediction (NCEP)/Department of Energy (DOE) Reanalysis 2 (R2) adjusts the top 10 cm soil moisture using the difference between model-derived and the 5-day mean of CPC Merged Analysis of Precipitation (CMAP) data \[Kanamitsu et al., 2002\]. However, this land assimilation system can reduce the quality of assimilation and thus the soil moisture predictability when the soil moisture analysis is not physically consistent with the atmospheric physics of the model. This is the case, for instance, of a heavy observed rain event producing a wet soil moisture analysis, while meanwhile the atmospheric physics of the model simulates an error such as a clear sky producing strong radiative and surface fluxes. Those strong surface fluxes are capable of reducing the predictability of the soil moisture analysis. Recent efforts have also been put into the NCEP Coupled Forecast System Reanalysis (CFSR) to produce a soil moisture
analysis as close to the truth as possible. CFSR performs uncoupled integration of its land surface model driven by the CMAP precipitation data every 24-hours. Then, the offline simulated soil moisture and soil temperature estimates are used as initial conditions of the CFSR for the following 24-hours. Since it is a similar offline land assimilation approach as used by GLACE-2, a spinup problem can be encountered (described above).

In this study, we use a physically consistent land assimilation system. It consists of assimilating 3-hourly precipitation observation-based data into the atmospheric component of a land-atmosphere coupled model by adjusting the vertical profile of the atmospheric humidity based on the difference between the model-derived and the observed precipitation. A Newtonian nudging of dynamical variables (surface pressure, vorticity, divergence, temperature) towards the R2 is also applied to reduce any model drift from the observed large scale atmospheric circulation. The combination of the dynamical nudging and the adjustment of the atmospheric humidity vertical profile not only brings the model-derived precipitation close to the observation but also redistributes the atmospheric heat and moisture, which in turn affects the adiabatic heating and hence the cloudiness. Then, unlike the R2, in this study, the radiative fluxes (directly affected by the cloudiness) and the surface fluxes are physically consistent with the soil moisture analysis.

The aim of this dissertation research is to produce a new soil moisture analysis and investigate its impact as initial conditions on the short-term (2 weeks) to subseasonal (2 months) forecasting skill of air temperature and precipitation. To our knowledge, the current state-of-the-art approaches in producing a soil moisture analysis as close to the truth as possible are either based on an offline land surface simulation driven by atmospheric forcing data (ex. GLACE-2) or on a land-atmosphere model in a coupled mode but with a non-physically consistent handling of the precipitation assimilating (ex. R2). As explained above, in these two approaches, the quality of assimilation can be impair, which in turn, can reduce the forecast skill. In this study, we use a physically consistent land assimilation system integrated into a land-atmosphere model in a coupled mode. It consists of assimilating a 3-hourly precipitation observation-based dataset into the atmospheric component of a coupled land-atmosphere model by adjusting the vertical profile of the atmospheric humidity based on the difference between the model-derived and the observed
precipitation. A Newtonian nudging of dynamical variables (surface pressure, vorticity, divergence, temperature) towards R2 is also applied to reduce any model drift from the observed large scale atmospheric circulation. The combination of the dynamical nudging and the adjustment of the atmospheric humidity vertical profile not only brings the model-derived precipitation close to the observation but also redistributes the atmospheric heat and moisture, which in turn affects the adiabatic heating and hence the cloudiness. Then, unlike the R2, the radiative fluxes (directly affected by the cloudiness) and the surface fluxes are physically consistent with the soil moisture analysis. The dissertation is organized as follows. Chapter 3 describes the soil moisture initialization system in details, the datasets used in this study and presents the results of the evaluation of the soil moisture analysis. Then, Chapter 4 presents the impact of this soil moisture analysis on the short-term to subseasonal forecasting skill of precipitation and air temperature fields. Finally, conclusions and prospectives of this work are in Chapter 5.
CHAPTER 3

DEVELOPMENT OF A SOIL MOISTURE ANALYSIS

The aim of this chapter is to describe and validate the soil moisture analysis that is used to initialize the forecasts in section 4. The initialization of global climate models with accurate soil moisture initial conditions (i.e. analysis) is a challenging task since there are no reliable global observations of soil moisture. The current state-of-the-art approaches for producing accurate soil moisture initial conditions are based on the assimilation of an observation-based atmospheric forcing data to drive the model-derived soil moisture estimates towards the truth as much as possible. However, to our knowledge, the approaches that have been used so far by the scientific community consist of producing either an offline land surface simulation (for example GLACE-2) or a simulation of a land-atmosphere model in a coupled mode but with a handling of the precipitation assimilation that is not physically consistent (for example NCEP R2 reanalysis). The offline assimilation approach may cause a spinup problem when producing the forecasts with the coupled land-atmosphere model since the climatology of the atmospheric forcing data and the coupled land-atmosphere are most likely to be different. The second approach used in the NCEP-R2 reanalysis responds to the spinup issue raised in the first approach by producing simulations in a coupled mode of a land-atmosphere model. However, because observationally-based precipitation are assimilated only at the land surface to adjust the top 10 cm soil moisture content, this approach is not always physically consistent. For instance, when the atmospheric component of the model simulates an error, such as producing a clear sky with consequently strong radiative and surface fluxes while a heavy observed precipitation event is assimilated into the land surface, in this case, the resulting wet soil moisture analysis is not physically consistent with the atmospheric physics of the model. Therefore, using these two state-of-the-art approaches can impair the quality
of assimilation and thus reduce the forecast skill.

In this study, we use a physically consistent assimilation approach using a land-atmosphere model in a coupled mode. The global land-atmosphere FSU/COAPS model used to assimilate the observation-based atmospheric forcing data and generate the summertime forecasts (analyzed in Chapter 4) is presented in section 3.1. The land assimilation technique used in this study is explained in detail in section 3.2. The experimental design and the soil moisture datasets used to evaluate and validate the soil moisture analysis are presented in section 3.4 and section 3.3 respectively. Finally, the validation results of the soil moisture analysis are presented in section 3.5.

### 3.1 Model

The FSU/COAPS climate model is a global spectral primitive equation model based on a Eulerian semi-implicit scheme \[\text{Cocke and LaRow, 2000}\]. The spatial resolution is a triangular truncation of 63 waves \((1.875^\circ \text{ latitude/longitude})\). The model uses weekly sea surface temperatures from \textit{Reynolds and al., 2002} at the boundaries. The land surface component of the model is the advanced National Center for Atmospheric Studies (NCAR)/Community Land Model \[\text{CLM2, Bonan et al., 2002}\]. Each grid cell in the CLM2 is represented by 5 primary subgrid land cover fractions (glacier, lake, wetland, urban, and vegetated). Each vegetated portion of the grid cell is divided into patches of up to 4 Plant Function Types (PFTs). The land surface component (CLM2) of the FSU/COAPS climate model produces prognostic soil moisture fields for 10 soil layers as opposed to the R2 with 2 soil moisture layers or the GSWP-2 with only one soil moisture layer. More details on the hydrological component of the CLM2 model are provided in Appendix A.

### 3.2 PAR Technique

As previously mentioned, there are no reliable global observations of soil moisture. Therefore, the land assimilation approach used in this study is based on the Precipitation Assimilation Reanalysis (PAR) technique. Precipitation is the main source of soil moisture and is expected
to affect its variability in depth. The PAR technique is similar to the physical initialization by Krishnamurti et al., [1984, 1988, 1991] but modified by Nunes and Cocke, [2004]. In order to reproduce the observed precipitation characteristics as much as possible, two nudging are performed. Dynamical variables (see below) are nudged towards R2 to avoid any drift of the model from observed large-scale circulation and precipitation is nudged towards observations (Figure 3.2.2). It is important to bear in mind that we do not assimilate the R2 precipitation because it is well known that while this reanalysis reproduces accurately the large-scale circulation, it has difficulties in reproducing the summertime precipitation. This may be because the climate model used to generate this reanalysis does not assimilate precipitation.

3.2.1 Precipitation Nudging

The precipitation nudging is defined by an analytic expression which modifies the humidity vertical profile as a function of the difference between model-derived and observed precipitation, in such a way that the model precipitation is brought closer to the observed precipitation. The analytic expression is a simple vertical structure function:

$$q_m = \frac{R_o}{R_p} q + \frac{(1/g) \int q d\sigma}{(1/g \int d\sigma)} \left( \frac{R_o - R_p}{R_p} \right)$$

Where $q_m$ is the modified vertical specific humidity profile, q is the specific humidity profile before PAR, and $R_o$ and $R_p$ are the observed and model-derived precipitation. The precipitation assimilation is not performed for precipitation rates less than 10 mm.d$^{-1}$.

3.2.2 Dynamical Nudging

The large-scale circulation is sensitive to the vertical distribution of the diabatic heating in the tropics. Since the precipitation assimilation modifies the vertical distribution of the diabatic heating, the model state can drift from observed large-scale circulation. Therefore, to reduce the model drift, prognostic variables (surface pressure, potential and virtual temperature, divergence and vorticity) are nudged toward the 6-hourly NCEP-R2 reanalysis that is thought to represent very well the observed large-scale circulation. The dynamical
nudging uses a Newtonian relaxation technique, which keeps the model variables close to the NCEP-R2 reanalysis by adding a nudging term in the prognostic equation, while still allowing the assimilation of precipitation. The Newtonian relaxation can be expressed as follows:

$$\frac{\partial \psi}{\partial t} = F(\psi) + \alpha(\psi - \psi_a)$$  \hspace{1cm} (3.2)

where $F(\psi)$ represents the FSU/COAPS climate model, $\alpha$ is the nudging term ($10^{-4}$ s$^{-1}$ for all dynamical variables), $\psi$ is the model-derived variable and $\psi_a$ is the variable from the NCEP-R2 reanalysis. The nudging term is applied at each model time step and is interpolated within the 6-hour interval. Since the PAR technique is based on a nudging technique, it is much computationally cheap.

![SOIL MOISTURE INITIALIZATION](image)

Figure 3.1: Schematic of the PAR technique.

### 3.2.3 Precipitation Datasets

Here, we describe the different precipitation datasets used for assimilation and validation of the PAR technique. In order to capture the diurnal cycle of the topsoil moisture state, a precipitation dataset at a time scale lower than daily is necessary for assimilation. Since there is no observation-based datasets with such a temporal scale lower than daily for the study period (1986-1995), we choose to use a bias-corrected reanalysis product. The study period
was selected to match with that of an international model intercomparison project named Global Land-Atmosphere Coupling Experiment (GLACE-2), of which the land-atmosphere FSU/COAPS model is part. Further explanation on the GLACE-2 experiment is given in Chapter 4.

3.2.3.1 Observations

Two types of observational precipitation dataset are available: rain gauge station measurements and satellite retrievals.

a. 3-hourly TRMM 3B42 (satellite retrieval)

Satellite retrievals are available most of the time globally and at a higher frequency than rain gauge measurements. For instance, the near-global Tropical Rainfall Measuring Mission (TRMM) 3B42 product is available at a very high temporal frequency (3-hourly) and a 0.25° spatial resolution [Huffman et al. 2003]. However, since the 1998-to-present time period covered by the TRMM 3B42 does not overlap the study period, this dataset is not selected for assimilation. Instead, we use a 3-hourly bias-corrected reanalysis product provided by Sheffield et al. [2006] described below (section 3.2.3.2 b). Nevertheless, the ultimate goal of this study is to use a fully observational dataset, such as the TRMM 3B42 data, to further improve the soil moisture initial conditions, which, in turn, may further increase the forecast skill.

The TRMM 3B42 product is based on the TRMM Multi-Satellite Precipitation Analysis (TMPA) provided by the National Aeronautics and Space Administration (NASA). The TMPA is derived by using an optimal combination of TRMM and other passive microwave precipitation estimates from instruments on board of different low-Earth-orbit satellites: the Special Sensor Microwave Imager (SSM/I), TMI, Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), and the Advanced Microwave Sounding Unit B (AMSU-B). In addition, information from one active instrument, the TRMM Precipitation Radar (PR), is merged into the TMPA dataset. In order to fill gaps left by the limited swath coverage of the low-Earth-orbit satellites, the merged microwave product is calibrated.
with infrared (IR) estimates from the rapid time capability of geostationary-Earth-orbit satellites [Huffman et al., 2006]. The IR radiometer brings the advantage of providing estimates with high spatial resolution and very good time sampling. Nevertheless, the IR radiometer measures the brightness temperature at the top of the cloud, which is only indirectly related to the surface rain rate. Then, a combination of both microwave and IR measurements provides the best estimate of precipitation. Finally, the dataset is scaled to match the monthly rain gauge analysis produced by the Global Precipitation Climatology Project (GPCP). This dataset also contains an estimate of the root mean square precipitation error. The error estimate is between 10% and 15% per grid point. Further details on this new satellite based precipitation dataset can be found in Huffman et al., [2006].

b. Monthly CRU and GPCC (rain gauge based datasets)

To validate the PAR technique over the land surface, we use a rain gauge based precipitation dataset. It is well known that satellite retrievals are very reliable over ocean but not over land surface because of the large variations of the surface emissivity caused by the rough topography of the land. In addition, the rain gauge based data offer a direct measure of rainfall and therefore are more accurate than the indirect measurements from satellite retrievals. Two following rain gauge based precipitation datasets are available for the study period: the Climatic Research Unit (CRU) and the Global Precipitation Climatology Centre (GPCC).

The CRU product is a 0.5° gridded dataset of monthly precipitation provides station data back to 1901 [New et al., 1999; 2000]. However, the gauge station database comprises less than 10,000 stations worldwide for the period 1986 to 2000. The CRU interpolates directly from station observations.

The GPCC product, operated by the German Weather Service, holds the largest rain gauge station database in the world with about 65,000-70,000 rain gauge stations worldwide for the period 1986-to-present, collected from the Global Telecommunications Network in real time, supplemented by other worldwide data collections, such as the Monthly Climatic Data for the World. The GPCC first interpolates available rain gauge data into a 0.5° grid,
then averaged them over a 2.5° grid [Rudolph et al., 1996]. Xie and Arkin, [1995] estimated the sampling error of the GPCC dataset and found that at least 5 gauges are required to be in a 2.5° grid box to keep the spatial sampling error within 10%. When only one gauge is available in a grid box, the sampling error is higher than 25%. The Figure 3.2 shows the number of gauges per 2.5° grid box. 83% of the total number of grid box over land contains at least 1 gauge station. In this study, the GPCC data are used to validate of the PAR technique, since this dataset comprises more rain gauge stations than CRU.

![Figure 3.2: Average number of gauges in 2.5° grid box per month between January 1986 and December 1995.](image)

**Figure 3.2:** Average number of gauges in 2.5° grid box per month between January 1986 and December 1995.

### 3.2.3.2 Bias-Corrected Reanalysis Products

As mentioned above, to capture the diurnal cycle of the topsoil moisture state, a precipitation dataset at a temporal scale lower than daily is necessary for assimilation. There is no such observational datasets for the study period. Therefore, we use a bias-corrected reanalysis product with a 3-hourly temporal frequency. Two 3-hourly, 1.0° bias-corrected reanalysis products are available for the study period:
a. 3-hourly GSWP-2 forcing data

The Second Global Soil Wetness (GSWP-2) observation-based atmospheric forcing dataset is based on the International Surface Climatology Project (ISLSCP-II) 3-hourly, 1.0° R2 reanalysis and scaled (i.e. bias correction) to the GPCC monthly data [Kanamitsu et al. 2002] and the CRU monthly data when GPCC is not available (described in section 3.2.3.1b).

b. 3-hourly Global Dataset of Meteorological Forcing (GDMF)

The GDMF product is constructed by combining the 6-hourly, 2.0° NCEP-National Center for Atmospheric Research (NCEP-NCAR) reanalysis with global observation-based datasets using a statistical downscaling in time and space [Sheffield et al. 2006]. Precipitation is spatially downscaled from 2.0° to 1.0° using the 1.0° Global Precipitation Climatology Project [GPCP, Huffman et al., 2001] daily dataset and temporally downscaled from 6-hourly to 3-hourly using the TRMM 3B42 (described above in section 3.2.3.1 a). The NCEP-NCAR reanalysis is known to show systematic biases at a monthly time scale. To remove these biases, the monthly totals of the NCEP-NCAR reanalysis are scaled to match that of the CRU monthly dataset. The reasons why we selected the GDMF over the GSWP-2 precipitation forcing data are the following: 1) through its spatial downscaling, the GDMF uses observed diurnal variability statistics based on TRMM observational data; in the GSWP-2 forcing data, the diurnal variability is from the R2 reanalysis and therefore much less reliable; an assimilation of precipitation data having an accurate diurnal cycle is very crucial for capturing the topsoil moisture diurnal cycle, which, in turn, affects the variability of the deeper soil moisture layers; 2) the GDMF adjusts the rain day frequencies to match observed statistics; this adjustment is not applied in the GSWP-2.

3.3 Soil Moisture Datasets

The different soil moisture datasets used to evaluate the soil moisture analysis obtained using the PAR technique are described in this section.
3.3.1 Global Soil Moisture Analysis Products

3.4.2.1 GSWP-2

Since there are no reliable global observations of soil moisture, an alternative to evaluating model soil moisture estimates on a global-scale is to use model analysis product. In this study, we use the Global Soil Wetness Project (GSWP-2) multi-model analysis integrating 13 offline state-of-the-art LSMs driven by the same atmospheric observation-based forcing. The forcing dataset used to drive the GSWP-2 multi-model analysis is described in section 3.2.3.2. Because it is the only available multi-model soil moisture analysis product using state-of-the-art LSMs driven by an atmospheric observation-based forcing for the study period, it is the best proxy for global soil moisture observations in this study. However, the GSWP-2 remains a model-based product and thus does not always ensure to be close to the truth since the state-of-the-art LSMs have not been fully validated. To estimate the surety of a GSWP-2 land surface variable, the standard deviation among LSMs of the global mean of the GSWP-2 outputs is provided [Dirmeyer, 2006]. One must keep in mind that, in this study, we use the GSWP-2 multi-model analysis to evaluate the model-derived soil moisture estimates on a global scale (i.e. as a proxy for global soil moisture observations) and not its atmospheric forcing dataset. The GSWP-2 is also used to evaluate other land hydrological estimates besides soil moisture, such as surface runoff and soil evaporation estimates.

3.4.2.2 NCEP-R2 Reanalysis (R2)

The performance of the PAR technique in producing a soil moisture analysis that is as close to the truth as possible, is compared in this study with that of R2. The land assimilation system used in R2 consists of correcting the top 10 cm soil moisture estimates using the difference between model-derived precipitation and observed pentad (5 day mean) precipitation [Kanamitsu et al., 2003]. This assimilation is applied solely at the land surface. Therefore, when the atmospheric physical processes (cloudiness and thus radiative fluxes) of their climate model are inconsistent with the soil moisture analysis, errors can be introduced and hamper the quality of assimilation. The R2 reanalysis uses the Oregon State University land surface model [OSU LSM, Pan and Mahrt 1987, Pan 1990], with two soil layers: a very thin top layer (0-10 cm) and a very thick deep layer (10 to 200 cm). The spatial resolution of R2 is T62 (1.915°) with 28 vertical levels.
3.3.2 In-Situ Observation Data

To evaluate the impact of the PAR technique on the associated soil moisture analysis, the best option is to use global soil moisture observational dataset provided, for instance, by satellite retrievals. However, despite significant advances, the current remote sensing techniques for soil moisture still suffer from issues associated with the shallow depth of the retrieval (less than 2 cm), the absence of retrieval over dense vegetated and frozen areas, and significant uncertainties in the retrieval algorithm. Only a few in-situ measurements of soil moisture sparse in time and space have been available.

![In-situ soil moisture stations (crosses) over Illinois. The red grid represents the grid points of the FSU/COAPS model.](image)

Figure 3.3: In-situ soil moisture stations (crosses) over Illinois. The red grid represents the grid points of the FSU/COAPS model.

Therefore, we choose to validate the impact of the PAR technique on the simulated soil moisture using in-situ soil moisture measurement from the Illinois Climate Network. This data network is provided by the Global Soil Moisture Data Bank [Robock et al. 2000] and comprises 19 stations covering the entire state (Figure 3.3). The soil moisture amount has been measured using neutron probes for 11 layers down to 2 m. A detailed description of
the dataset and its measurement errors are given in *Hollinger and Isard*, [1994].

### 3.4 Experimental Design

Two global numerical simulations are carried out from January 1986 to December 1995 (Figure 3.4). For the first simulation, using the PAR technique described in the previous section (3.2), the 3-hourly observed precipitation is continuously assimilated (hereafter, PAR simulation). The 2nd simulation is performed without assimilation (hereafter, CONTROL simulation). In other words, the CONTROL simulation is a free run.

![CONTROL SIMULATION](image)

![PAR SIMULATION](image)

**Figure 3.4:** Two numerical simulations using the land-atmosphere FSU/COAPS model.

The soil moisture evolves slowly. Therefore, land models require a long spin-up time period to reach equilibrium. Using the Community Land Model Version 3 (CLM3), *Du et al.* [2006] found that the equilibration state of soil moisture at 1.5 m depth is achieved after at least 20 years. In the FSU/COAPS climate model, the total soil moisture depth goes down to 3.4 m. We found that the global average soil moisture of the deepest layer reaches an equilibrium state at the end of a 50-year run (Figure 3.5). Both simulations start after being spun-up.
Figure 3.5: Global soil moisture average of the deepest layer (229 to 343 soil depth) during the last 8 years of the 50-year spinup time period.

3.5 Results

3.5.1 Precipitation Assimilation Verification

Before analyzing the impact of the PAR technique on the model-derived land surface hydrological estimates (in particular soil moisture), we first verify whether the land-atmosphere FSU/COAPS model is able to assimilate the GDMF precipitation dataset (described earlier in section 3.2) over the land surface. To assess the ability of the model to reproduce observed precipitation patterns regardless of its magnitude, spatial and temporal correlations are computed. Another commonly used statistical tool for precipitation validation is the Equitable Threat Score (ETS) in conjunction with the bias score [Schaefer, 1990]. In contrast with the correlation, the ETS score takes into account the spatial distribution and the magnitude of precipitation. The ETS score is primarily the ratio of the number of hits to the sum of forecast hits and misses at a given precipitation threshold. A hit is defined as a grid point where both the simulated and observed precipitation are equal to or exceed the given precipitation threshold. The bias score reveals the systematic model error at the given precipitation threshold. A bias greater (lower) than 1 indicates that the model precipitation is over- (under-) estimated. Finally, a perfect score is equivalent to ETS = 1 and bias = 1. In the computation of both correlations, ETS and bias score, all continental grid points are
Figure 3.6 represents the average continental precipitation over the boreal summer months (June, July, August) and the boreal winter months (December, January, February) between 1986 and 1995. For both seasons, considerable differences in the spatial distribution are noted between CONTROL and GDMF. PAR and R2 appear to be in very good agreement with GDMF; both model estimates depict very well the belt of maximum precipitation (South America, Central Africa and Southeast Asia) associated with the seasonal ITCZ path.

Figure 3.7 shows the results of the 4 statistical tools calculated against GDMF on a daily time scale between 1986 and 1995 at 10 different precipitation thresholds. For both seasons, PAR exhibits distinctly better statistical results than CONTROL or R2. The temporal and spatial correlations between PAR and GDMF are both superior to 0.7 and the ETS values are superior to 0.4 at all thresholds except at low precipitation thresholds. The comparison of the two seasons shows that PAR gives better results in the boreal summer than the boreal winter. This is expected since the PAR technique essentially modifies the convective precipitation occurring during the summer. The high spatial correlations between R2 and GDMF for both season (> 0.7) confirm the result found earlier in Figure 3.6 that R2 well reproduces the GDMF spatial distribution. However, R2 poorly agrees in terms of temporal variability (correlation < 0.3). Finally, for both seasons, the bias score indicates that the three model estimates have a tendency to underestimate the high precipitation amounts and overestimate the low precipitation amounts of the GDMF dataset.

Since the GDMF precipitation dataset has not been validated against an observational dataset, it is worth evaluating the PAR simulation against an independent observational precipitation dataset. We choose to evaluate the PAR precipitation against the GPCC observational data available at monthly scale because no gridded observations at a daily time scale are available for the study period (Figure 3.8). Figure 3.8 shows similar results as Figure 3.7. For both seasons, PAR exhibits higher spatial and temporal correlations and higher ETS values for most thresholds than CONTROL and R2. Nevertheless, R2 shows significantly higher spatial and temporal correlations and ETS scores at a monthly than a daily scale. This may indicate that R2 is significantly more reliable at a monthly than a daily time scale. In addition, R2 is closer to observations in the boreal winter than in the
boreal summer, which is a well-known feature of R2 at a monthly scale. In fact, in the boreal winter, the R2 obtains even higher ETS values than the PAR simulation for low precipitation thresholds.

**Figure 3.6:** Average precipitation (mm/month) over the a) boreal summer months and b) boreal winter months from 1986 to 1995.
a) Daily Boreal Summer (JJA)  
b) Daily Boreal Winter (DJF)

Figure 3.7: Statistical tools calculated against the daily totals of GDMF for the a) boreal summer (JJA) and b) boreal winter (DJF). All grid points of the global scale are taken into account.
Figure 3.8: Same as Figure 3.7 but against the monthly GPCC dataset.
3.5.2 Local Land Surface Variables Evaluation

3.5.2.1 Soil Moisture Validation

Based on the high correlation values (> 0.9 in boreal summer and > 0.8 in boreal winter) between the model-derived precipitation estimates and two precipitation datasets (the GDMF precipitation forcing data and an independent precipitation dataset), we have shown in the previous section that, the land-atmosphere FSU/COAPS model is able to assimilate the GDMF precipitation data and to reproduce the observed spatial and temporal precipitation variability. Using in-situ measurements over Illinois, in this section we locally validate the impact of the PAR technique on model-derived soil moisture estimates by analyzing its seasonal cycle and temporal anomaly characteristics. The soil moisture analysis is also compared to that of the R2, which also uses a land assimilation system produced in a coupled mode of a land-atmosphere model. The metric used to assess which soil moisture analysis is closer to observations is the anomaly correlation calculated against in-situ observations.

a. Climatology

Figure 3.9 shows the seasonal cycle of the upper 10 cm soil moisture and the deeper soil moisture (10-200 cm) at 6 different locations over Illinois for the period 1986-1995. The different locations correspond to the 6 model grid cells inside the state of Illinois and the stations inside each model grid cell are aggregated (as described in Figure 3.3). The bars represent the standard deviation among the soil moisture stations. Note that grid cells 3 and 6 do not possess bars because they correspond with only one station. In the topsoil layer and at most of the locations, CONTROL reproduces well the sharp decrease of soil moisture in the summer but is out of phase in the winter. Despite its weak amplitude and dry bias, the PAR simulation follows the observed seasonal cycle best. R2 has a nearly constant soil moisture estimate at all locations and therefore fails to capture the observed soil moisture seasonal cycle in the topsoil layer. This issue may be related to the simplified freezing process used in R2, which assumes that when the air temperature drops to the freezing temperature, precipitation and melted snow cannot infiltrate the soil [Boisserie et al. 2006; Li et al. 2005]. Thus, the soil moisture in R2 is not capable of recharging during cold winters. In contrast, R2 follows surprisingly well the observed seasonal cycle in the deep soil layer (Figure 3.9b).
R2 presents thus a physical inconsistency between its top 10 cm and the deep soil moisture layer. We speculate that this physical inconsistency in R2 can be due the combination of the two following factors: 1) its precipitation forcing is applied only at the top 10 cm soil moisture layer; and 2) its deep soil layer is very thick compared to the top layer, which favors the prolongation of the adjustment time of the deep soil moisture layer to the top layer forcing. The behavior of the two numerical simulations in the deep soil layer is similar to that of the top soil layer (Figure 3.9b). Indeed, at all locations, the CONTROL is again out of phase in the boreal winter and the PAR simulation follows the observed seasonal cycle best. Finally, it is clear that, in both soil layers, CONTROL and PAR have a systematic dry soil moisture bias. This model error may be attributed to an incorrect partitioning of the evapotranspiration in the CLM2 model (the land surface component of the land-atmosphere FSU/COAPS model) as highlighted in a paper by Lawrence et al. [2007]. In this latter paper, it was found that several modifications of parameterization, such as increasing transpiration and infiltration and, decreasing soil evaporation, greatly reduce the dry soil moisture bias and increase the soil moisture seasonal cycle amplitude. These modifications lead to a new version of the community land model (CLM3) that we hope to use in the near future.
Figure 3.9: Monthly soil moisture climatology of the in-situ observation (thick grey), CONTROL (dotted), PAR (dashed), R2 (semi-dotted) and GSWP2 (thin grey) in a) the top 10 cm layer and b) the deeper 10-200 cm layer at 6 locations in Illinois.
Here, we validate the soil moisture anomalies up to 2 m depth. Figures 3.10 and 3.11 display the vertical profile of the average soil moisture anomalies of in-situ observations, R2, PAR and CONTROL over Illinois for 1986-1990 and 1991-1995 respectively. PAR, CONTROL and the observations are discretized into the same 7 soil layers (0-10, 10-30, 30-50, 50-70, 70-90, 90-110, 110-200 cm), while the R2 has only two soil layers (0-10 and 10-200 cm). First, it is clear that the R2 shows a large discontinuity in soil moisture anomalies between its two layers, which may be due to the combination of two factors: 1) the precipitation forcing is only applied at the top layer; and 2) the thick deep layer prolongs the response time to the top layer. PAR and R2 reproduce reasonably well the major dry and wet events occurring in the state of Illinois, such as the 1988 drought and the 1993 flood. However, one can notice that the anomaly amplitude in PAR is most of the time too weak. As previously mentioned, the land surface component of the land-atmosphere FSU/COAPS model has an incorrect evapotranspiration partitioning and produces too much soil evaporation within the canopy. Our speculation is that strong soil evaporation is likely to weaken the soil moisture response to precipitation and hence weaken the amplitude of soil moisture anomalies. CONTROL picks the signals of the 1993 flood but fails in reproducing the 1988 drought. CONTROL is even sometimes out-of-phase with observations, such as during the summer of 1995. The anomaly correlations computed at the same locations defined in Figure 3.3 show that, in both soil layers, PAR correlates the best with the observations at all 6 locations (Figure 3.14). These results indicate that the PAR technique is: 1) able to greatly improve the model-derived soil moisture anomalies in the land-atmosphere FSU/COAPS model, and 2) these anomalies are even closer to observations to those of the R2.

Because the GSWP-2 analysis product is latter used to evaluate soil moisture estimates on the global-scale, here we want to compare its values with the in-situ observations since this product will be used as a proxy for global soil moisture observations in section 3.5.3. Figure 3.13a shows the average soil moisture seasonal cycle of the two numerical simulations, R2 and GSWP-2 over Illinois. For comparison purposes, all model-derived soil moisture estimates are scaled to the 1.5 m soil moisture depth of GSWP-2. Like the two numerical simulations, the GSWP-2 analysis shows a dry bias. Figure 3.13b shows that GSWP-2 follows
very well the observed seasonal cycle and amplitude. Based on the high anomaly correlation (0.85), the GSWP-2 also follows the observed interannual variability (Figure 3.13b). The PAR simulation also shows a high correlation with observations (0.78) but lower than that of the GSWP-2. This could be due to the smoothing of the GSWP-2 multi-model analysis since it represents the average of 13 models.

Figure 3.10: Vertical profile of soil moisture anomalies for the period 1986-1990.
Figure 3.11: Vertical profile of soil moisture anomalies for the period 1991-1995.

Figure 3.12: Monthly soil moisture anomaly correlation a) in the top 10 cm layer and b) in the deep layer 10-200 cm at 6 different locations in Illinois.
Figure 3.13: Monthly soil moisture a) seasonal average and b) anomaly correlation against the in-situ observation with respect to 1986-1995 time period in the top 1.5 cm layer averaged over Illinois. For the plot of seasonal average, the thick grey line indicates the in-situ observations, the dotted line indicates the CONTROL simulation, the dashed line indicates the PAR simulation, the semi-dotted line indicates R2 and the thin black line indicates GSWP2.

3.5.2.2 Land Surface Water Budget Components Evaluation

The PAR technique not only affects the model-derived soil moisture estimates but also the estimates of other components of the surface water budget, such as surface runoff and soil evaporation. The evaluation of these two variables is a difficult task due to the lack of observations. Here, we evaluate these variables against the GSWP-2 multi-model analysis product by comparing the seasonal cycle and the anomaly correlation over Illinois. Since the GSWP-2 analysis is the mean of several LSMs, the bars represent the variance among LSMs as a measure of uncertainty of this product. However, one must be cautious when interpreting these bars. A small variance does not necessarily mean that the GSWP-2 is close to observations. However, a large variance is a good indicator that the GSWP-2 is not reliable.

Figure 3.12 presents the average seasonal cycle, the standard deviation and the anomaly correlation of precipitation, surface runoff and soil evaporation over Illinois for the study period. For precipitation, PAR agrees, not surprisingly, very well with GDMF and exhibits the best anomaly correlation (0.82). These two results simply corroborate the success of the GDMF precipitation assimilation into the land-atmosphere FSU/COAPS model found earlier in section 3.5.1. R2 also does a good job in following the observed seasonal cycle and
obtains a reasonably high anomaly correlation (0.67). CONTROL shows poor agreements with GDMF, with an overestimation during the spring and fall and a very low anomaly correlation (< 0.2).

The comparison of the surface runoff estimates shows that all three datasets, CONTROL, PAR and R-2 have difficulties in following the seasonal cycle of GSWP-2. PAR seems to be the least in agreement with GSWP-2 with a weak seasonal cycle amplitude and a dry bias. In contrast, the PAR surface runoff shows a significant increase of anomaly correlation over CONTROL (increase from 0.05 to 0.6). The PAR anomaly correlation is even higher than that of R-2 (0.42). Nevertheless, note that GSWP-2 has large error bars. This indicates that the different LSMs contributing to the GSWP-2 do not converge to the same surface runoff analysis and therefore the reliability of the GSWP-2 to estimate surface runoff state is small. The finding of a high disparity among models for surface runoff in GSWP-2 is consistent with the paper by Dirmeyer et al. [2005] who computed the standard deviation of the global mean for each GSWP-2 land surface analysis.

Finally, all the datasets have a soil evaporation maximum value during the warm season. However, the summer peak of PAR and the R-2, occurring in July, best match that of GSWP-2, while the summer peak of CONTROL is a month early. The relatively small bars indicate that GSWP-2 soil evaporation estimate is most likely reliable. PAR obtains the best anomaly correlation (0.51) compared to -0.08 for CONTROL and 0.48 for R-2. To conclude, the PAR technique not only improves the model-derived soil moisture estimates but also the estimates of surface runoff and soil evaporation over Illinois. The comparison with R-2 shows that these estimates are most of the time closer to observations that these of R-2. However, a local validation over Illinois is not sufficient to conclude that the PAR technique is capable of generating realistic soil moisture initial conditions for the global FSU/COAPS climate model.
Figure 3.14: Climatology (1st column), standard deviation (2nd column) and anomaly correlation of the average monthly a) precipitation, b) surface runoff and c) soil evaporation across Illinois (1986-1995).
3.5.3 Global Land Surface Variables Evaluation

Here, we evaluate the effect of the PAR technique on a global scale against a proxy for global land surface observations, the GSWP-2 multimodel analysis (described in section 3.4.2.1). To evaluate the model-derived soil moisture, surface runoff and soil evaporation estimates, we analyze the global amount averaged the summer and winter seasons and the time anomaly correlations with respect to GSWP-2. Since GSWP-2 possesses only one soil moisture layer (top to 1.5 m), the soil moisture layers of PAR and CONTROL are averaged to match the GSWP-2 soil depth. For R2, which has two soil layers (0-10 and 10-200 cm) and we scale its soil moisture estimate to 1.5 m depth as follows [Li et al. 2005]:

\[ SM = 1500(0.10 \times SM_1 + 0.90 \times SM_2) \]  \hspace{1cm} (3.3)

Where \( SM \) represents the total soil moisture (mm) in the top 1.5 m; \( SM_1 \) is the volumetric soil moisture in the top 10 cm layer and \( SM_2 \) is the volumetric soil moisture in the 10-200 cm layer.

Figure 3.15 shows the global soil moisture storage for the upper 1.5 m of soil for both summer and winter seasons. During both seasons, all soil moisture estimates capture the large-scale climate patterns of the earth (dry deserts and wet rain forests). However, CONTROL underestimates the soil moisture amount in most of the wet regions and misplaces the maxima. PAR and the R2 appear to reproduce well the small-scale features and amplitude of the GSWP-2 soil moisture estimate, such as the wet regions in the path of the ITCZ (the northern part of South America, Central Africa and Southeast Asia). One also can notice that while PAR agrees well with the GSWP-2 soil moisture absolute amount, the R2 seems to overestimate it in most regions of the globe.

For both seasons, all analyses capture the large-scale climate features of the GSWP-2 surface runoff estimates (Figure 3.16). They also all agree that Asia is wetter in the summer than in the winter due to the summer Asian monsoon. However, it is clear that R2 and PAR reproduce better the spatial distribution of GSWP-2 than CONTROL for both seasons.

Although R2 captures well the seasonal large-scale pattern of soil evaporation, it over-
estimates the soil evaporation amount of GSWP-2 (Figure 3.17). Since the soil evaporation amount is directly related to the available soil moisture amount, this result can be explained by the overestimation of soil moisture already seen in Figure 3.15. PAR compares well with GSWP-2 in terms of spatial distribution and amplitude.

Finally, Figure 3.18 shows the spatial distribution of temporal anomaly correlations with respect to GSWP-2 for each land surface hydrological estimates (soil moisture, surface runoff and soil evaporation). Between -0.4 and 0.4, the temporal anomaly correlation is not statistically significant at the 99% confidence interval. For soil moisture, it is first obvious that the temporal anomaly correlation of PAR is significantly increased over CONTROL. In comparison with R2, PAR correlates slightly better in most regions (South America, Africa, Southeast Asia and Australia), except in the very high latitudes of the Northern Hemisphere. For surface runoff correlation, PAR shows clearly higher temporal anomaly correlations than CONTROL and R2 across the globe. The evaporation time anomaly correlation is very high across the land surface for both PAR and R2, except over the tropical rainforests (the Amazon and Central Africa).

To conclude, it is found that the land surface hydrological variables, in particular soil moisture analysis compares well with GSWP-2 across the globe.
Figure 3.15: Average soil moisture (mm/month) from 1986 to 1995 of a) the boreal summer months and b) the boreal winter months.
Figure 3.16: Average surface runoff (mm/month) from 1986 to 1995 of a) the boreal summer months and b) the boreal winter months.
Figure 3.17: Average surface evaporation (mm/month) from 1986 to 1995 of a) the boreal summer months and b) the boreal winter months.
3.5.4 Land Surface Water Budget

Unlike a free run, a simulation with precipitation assimilation (such as PAR) can lead to water mass imbalance. In order to investigate this closure problem, here we examine the land surface water budget of the land-atmosphere FSU/COAPS model and compare it with that of GSWP-2 and R2. The land surface water budget is expressed as follows:

Figure 3.18: Time anomaly correlation between each numerical simulation and observations for a) soil moisture b) surface runoff and c) surface evaporation across the years from 1986 and 1995.
\[ N = P - E - R - SM \] (3.4)

where P is the precipitation, E the evaporation, R the surface runoff, SM the soil moisture and N the residual term (i.e. non-closure term). Figure 3.19 shows the annual mean of the residual term (N) across all months from 1986 to 1995. It is shown that the residual term of PAR is not significantly greater than that of CONTROL. This indicates that the PAR technique does not lead to a surface water imbalance. It is clear that R2 shows an excess of water (N<0) in the land surface across the entire globe, which is most likely because the precipitation forcing in R2 is applied only at the land surface to adjust soil moisture level [Lu et al., 2005]. This may suggest that the PAR technique is a more physically consistent land assimilation technique than that of R2. Finally, CONTROL, PAR and GSWP-2 show a slight excess of water in the atmosphere (N>0) that could be due to the fact that the above surface water budget (equation 3.4) does not take the melting snow and subsurface runoff, which are both sink terms for the atmosphere water storage, into account.

Figure 3.19: Surface water budget as defined in equation 3.4 of a) the GSWP-2, b) the R2, c) the PAR and d) the CONTROL average from 1986 to 1995.
3.6 Summary and Discussion

In this chapter, we have evaluated the impact of a physically consistent land assimilation technique on model derived land surface estimates in the land-atmosphere FSU/COAPS model. The evaluation focuses mostly on the soil moisture estimates since it is a key land surface variable for having accurate forecasts of near surface variables. First, based on temporal correlations (> 0.8) and spatial correlations (> 0.8) between the model-derived and GDMF precipitation data in boreal and winter, we found that, the land-atmosphere FSU/COAPS model is able to assimilate the GDMF precipitation dataset. In addition, the model derived precipitation estimates are in better agreement with GDMF and an independent precipitation observational dataset (GPCC) than the R2 reanalysis. Then, the impact of this land assimilation technique has shown overall a positive impact on soil moisture estimates over Illinois as far as the seasonal cycle variability, its amplitude and the anomaly variability (anomaly correlations > 0.5). We also found that this assimilation technique not only improves soil moisture estimates but also other hydrological land surface variable, such as surface runoff and soil evaporation estimates. The comparison with a proxy for global land surface observations (GSWP-2 multi-model analysis) suggests that the land assimilation system is effective not only over Illinois but across the globe. In boreal summer and winter, the spatial distribution and amplitude of the soil moisture, surface runoff and soil evaporation estimates compare better with GSWP-2 than the free run without precipitation assimilation (i.e. CONTROL simulation) and the R2 reanalysis. The soil moisture anomaly correlations with respect to GSWP-2 are also higher (coefficients up to 0.8 across most of the globe) than those of the free run without assimilation (coefficients lower than 0.4 across most of the globe). Compared to the R2 reanalysis, the anomaly correlations are slightly higher. However, one should be cautious in interpreting those results. Although the GSWP-2 offers the best proxy for land surface observations because it represents the average of 13 state-of-the-art LSMs, it is not necessary always accurate. Indeed, since these state-of-the-art LSMs have not been fully validated, they cannot ensure to be close to the truth. In particular, GSWP-2 is not most likely to give an accurate estimates when there is a high variability among these LSMs, which is the case for the GSWP-2 surface runoff estimate. Finally, the finding that overall the soil moisture analysis produced in this study better compares with in-situ observations and a proxy for global observations across the globe than the R2
reanalysis may suggest that a physically consistent land assimilation technique is important for generating accurate soil moisture initial conditions.

To conclude, in this section, we have developed a new soil moisture analysis product that is: 1) physically consistent with the atmospheric physics of the coupled land-atmosphere FSU/COAPS model; 2) close to observations over Illinois and; 3) comparable to a benchmark in soil moisture analysis (the GSWP-2 multi-model analysis). Therefore, this soil moisture analysis dataset can be used in many land surface studies, such as crop modeling, detection of extreme events (drought and flood), water management. In the next section, this dataset is used to initialize short-term to subseasonal forecasts of near-surface air temperature and precipitation.
CHAPTER 4

THE IMPACT OF A NEW SOIL MOISTURE ANALYSIS ON SHORT-TERM TO SUBSEASONAL FORECASTING SKILL

In the previous chapter, we produce a soil moisture analysis using the PAR technique. The aim of this chapter is to investigate the impact of this analysis on the short-term to subseasonal forecasting skill of summertime 2-m air temperature and precipitation. Given the long soil moisture memory (i.e. anomaly persistence) and knowing the effect of soil moisture fields on both the surface energy and water budget, it is thought that, during the boreal summer, accurate soil moisture initial conditions can potentially increase the subseasonal forecast skill of near surface variables.

4.1 GLACE-2 Overview

The second phase of the Global Land-Atmosphere Coupling Experiment (GLACE-2) is an international ongoing project that has similar goals as this study. It aims at improving the subseasonal forecasts of precipitation and air temperature through a realistic initialization of soil moisture among a wide variety of climate models. This model intercomparison project is a follow-up to the GLACE-1 project, which sought to quantify the degree to which simulated precipitation responds to prescribed times series of soil moisture content [Koster et al., 2006]. The results of GLACE-1 helped to identify the regions with a relatively high land-atmosphere coupling strength. These regions are named "hotspots" and are displayed in Figure 4.1:

As part of my Ph.D research, I have joined this international research group using the land-atmosphere FSU/COAPS model (described in section 3.1). The participation of the
Figure 4.1: Land-atmosphere coupling from GLACE-1 by Koster et al., [2004]. The black boxes are the so-called "hotspots", regions where coupling strength is high.

Land-atmosphere FSU/COAPS model into the GLACE-2 experiment presents the unique opportunity to compare the results of the our model with the other climate models listed in Table 4.1. It also provides an environment in which COAPS collaborates with other research groups to better understand the climate variability.

As explained in the previous chapter (section 3.2), the FSU/COAPS model initializes its forecasts with the PAR technique, which consists of assimilating precipitation and dynamical variables throughout an online integration of the model (i.e. coupled with the land surface model). This land assimilation technique is different from the offline assimilation technique (described in section 3.3) used by almost all participants in GLACE-2. Almost all participants in GLACE-2 drive their land surface models off-line with the GSWP-2 observation-based atmospheric forcing data from 1986 to 1995. The land surface state variables from their offline simulation are then used to initialize their coupled land-
atmosphere model (i.e. climate models). However, a common problem is encountered when using an offline assimilation technique is that it is well known to each model has its own climatology, which is also different from that observed. This means that the climate models used by the participants in GLACE-2 are biased with respect to the GSWP-2 observation-based atmospheric forcing data. This bias problem results in prolonging their climate model spinup time, which then can decrease the short-term to subseasonal forecast skill of near surface variables. To reduce the effect of the spinup problem, each participant of GLACE-2 adjusts their land initial fields \(X_{\text{offline}}\) to their climate model climatology \(X_{\text{online}}\) before using them for initializing the forecasts as follows (equation 4.1):

\[
X_{\text{online}} = (\frac{X_{\text{offline}} - \bar{X}_{\text{offline}}}{\sigma_{X_{\text{offline}}}})X_{\text{online}} + \bar{X}_{\text{online}}
\] (4.1)

Where \(X\) is the timeseries of a land surface variable (for example, soil moisture) spanning the years 1986 to 1995. The bar denotes a temporal average of \(X\) and \(\sigma_X\) represents the standard deviation of \(X\) over the 10-year study period. The problem with this adjustment is that it is not physically based. It is a simple mean bias correction. In the land assimilation technique used in this study (described in section 3.2), this adjustment is not required because the land initial fields and the forecasts are generated using the same model (the coupled land-atmosphere FSU/COAPS model). This implies that any bias is minimized since the initial fields have the same climatology as that of the forecasts.

In the land assimilation technique used in this study (described in section 3.2), any bias problem is minimized because the land initial fields and the forecasts are generated using the same model (the coupled land-atmosphere FSU/COAPS model). In this case, the initial fields have the same climatology as that of the climate model. Therefore, this adjustment used in GLACE-2 is not required in the FSU/COAPS climate model.

Some of the unpublished results of the GLACE-2 will be displayed and briefly compared with the results from the FSU/COAPS climate model in section 4.4.3.
Table 4.1: List of participants in the GLACE-2 experiment.

<table>
<thead>
<tr>
<th>Groups: Models</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center for Ocean-Land-Atmosphere Studies (COLA) GCM V3.2</td>
<td>Misra et al. [2007]</td>
</tr>
<tr>
<td>European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System</td>
<td>Bechtold et al. [2008]</td>
</tr>
<tr>
<td>NASA/Global Modeling and Assimilation Office (GMAO) seasonal forecast system (pre-GEOS5 version)</td>
<td>Bacmeister et al. [2000]</td>
</tr>
<tr>
<td>National Center for Atmospheric Research (NCAR) Community Atmospheric Model 3.0</td>
<td>Collins et al. [2006]</td>
</tr>
<tr>
<td>National Center for Atmospheric Research (NCAR) Community Atmospheric Model 3.5</td>
<td>Neale et al. [2008]</td>
</tr>
<tr>
<td>Community Land Model 3.5</td>
<td>Oleson et al. [2008]</td>
</tr>
<tr>
<td>FSU/COAPS</td>
<td>Shin et al. [2005]</td>
</tr>
<tr>
<td>Geophysical Fluid Dynamics Laboratory (GFDL) Global Atmospheric Model</td>
<td>Cocke et al. [2000]</td>
</tr>
<tr>
<td>National Centers of Environmental Prediction (NCEP) Global Forecast System (GFS/Noah)</td>
<td>GFDL Team [2004]</td>
</tr>
<tr>
<td>Royal Netherlands Meteorological Institute (KNMI): European Centre for Medium-Range Weather Forecasts (ECMWF)</td>
<td>Delworth et al. [2006]</td>
</tr>
<tr>
<td></td>
<td>Moorthi et al. [2001]</td>
</tr>
<tr>
<td></td>
<td>Kalney et al. [1996]</td>
</tr>
<tr>
<td></td>
<td>Pan and Mahrt [1987]</td>
</tr>
<tr>
<td></td>
<td>Ek et al. [2003]</td>
</tr>
<tr>
<td></td>
<td>Van der Hurk and P. Viterbo [2003]</td>
</tr>
</tbody>
</table>

4.2 Experimental Design

4.2.1 Description

We follow the same experimental design as in the GLACE-2 experiment to generate forecasts. This experimental design consists of running two series of 2-month retrospective forecasts (i.e. hindcasts). The first series of forecasts (hereafter, PAR forecasts) is initialized with the land surface analysis produced using the PAR technique (described in Chapter 3). The second series of forecasts (hereafter, control forecasts) is initialized with land surface values.
obtained from a free run (i.e. non-assimilated run). Table 4.2 summarizes the experimental design. Because we aim to evaluate the model response to the land surface initialization, we want to isolate the influence of atmospheric initial conditions and ocean boundary conditions. For this purpose, sea surface temperatures (SSTs) are set to weekly observed values [Reynolds and al., 2002] and the atmospheric initial conditions are taken from the R2 reanalysis for both series. Thus, by subtracting the two series, the influence of atmospheric initial conditions and ocean boundary conditions are strongly reduced over that of the land surface initial conditions. Both series consist of 60 independent forecasts; one for each of the six starting dates (June 1, June 15, July 1, ... August 15) in each of the ten years spanning 1986 to 1995.

For each of the 60 forecasts, 10 ensemble members are generated. The ensemble members are produced using different atmospheric initial conditions from the R2 reanalysis every 6 hours before the forecast starting time. For instance, the 10 ensemble members of the forecast starting in 1 June 1986 are generated using the atmospheric initial conditions taken at (1 June 00h, 31 May 18h, 31 May 12h,..., 29 May 18h). We take the ensemble mean over all individual ensemble members because by taking the mean, we are able to remove most of the model internal variability and therefore we have a better representation of the climate variability.

Table 4.2: Summary of the experiments: name of the series of forecasts, its description, number of forecasts, study period, time scales of the forecasts, and equivalent name used in the GLACE-2 experiment.

<table>
<thead>
<tr>
<th>Series</th>
<th>PAR forecasts</th>
<th>Control forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>initialized with the PAR technique</td>
<td>initialized with a free run</td>
</tr>
<tr>
<td>Number of forecasts</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Time scales</td>
<td>15-days, 1-month, 2-month</td>
<td>15-days, 1-month, 2-month</td>
</tr>
<tr>
<td>Name of series in GLACE-2</td>
<td>SERIES-1</td>
<td>SERIES-2</td>
</tr>
</tbody>
</table>

As already mentioned, through its control of the partitioning into sensible and latent heat fluxes, soil moisture is believed to affect many near surface variables. In this study, we analyze the forecasts of two near surface variables, precipitation and 2-m air temperature. While the model-derived precipitation estimate is calculated in the atmospheric component of
the land-atmosphere FSU/COAPS model, the model-derived 2-m air temperature estimate is calculated in the land surface component (i.e. CLM2 described in section 3.1). The calculation of the 2-m air temperature estimate is performed for two different types of soil, bare soil and soil beneath canopy.

\[
t_{2m} = t_1 + \frac{\delta tv}{k} \log\left(\frac{2 + z_0}{z_0}\right)
\] (4.2)

where \(t_{2m}\) is the 2-m air temperature, the relation for potential temperature, \(\delta tv\) is the difference of virtual temperature, \(k\) is the von Karman constant (\(k=0.41\)) and \(z_0\) is the roughness length of sensible heat.

\[
\begin{align*}
  t_1 &= \text{tg} \; \text{(if bare soil)} \\
  t_1 &= \text{taf} \; \text{(if soil beneath canopy)}
\end{align*}
\]

Where \(\text{tg}\) is the ground temperature and \(\text{taf}\) is the air temperature within canopy.

### 4.2.2 Metrics To Measure Forecast Skill

Because the forecast starting dates are every two weeks, we are able to construct a time series using all 60 forecasts and calculate the time anomaly correlation between the forecasts and the observations at each grid point. This is the metric used here to evaluate the forecast skill of each series of forecasts. Then, to measure the increase (or decrease) in skill attributed to the soil moisture analysis, we calculate the difference in time anomaly correlation squared \((r^2)\) between the PAR forecasts and control forecasts. We choose to calculate the difference in \(r^2\) instead of the difference in \(r\) because the subtraction of two correlations is not always meaningful. On the contrary, the difference in \(r^2\) is meaningful and indicates the increase of the fraction of explained variance attributed to the soil moisture analysis. Nevertheless, there are two cases in which one needs to be cautious. These cases are when one of the correlations is negative or not statistically significant. To overcome this problem, the difference in \(r^2\) is set to zero in these two cases. In this chapter, we also analyze the forecast for two particular years during which extreme hydrological events occurred, 1988 and 1993.
4.3 Observational Datasets

Here, we describe the datasets used for the validation of the forecasts. The forecasts are analyzed at three different time scales: first 15 days average, first month average and two-month average (Table 4.2). The observational dataset used for the evaluation of the precipitation forecast is provided by Higgins et al., [2000]. The data are daily precipitation on a 0.25° latitude/longitude grid over the continental United States from 1948 to 1998 and defined by interpolating quality controlled gauge observations at over 8000 stations collected from multiple sources.

To evaluate the 2-m air temperature forecasts we use two different datasets. Global data sets of 2-m air temperature gridded observational data at scales less than a month are not available. Therefore, for the short-term forecasts (15-day time scale), we use a bias-corrected reanalysis product. This product is the Global Dataset of Meteorological Forcing (GDMF) provided by Sheffield et al., [2006]. It is constructed by combining the 6-hourly 2.0° NCEP-National Center for Atmospheric Research (NCEP-NCAR) reanalysis with the most recent global observation-based datasets using a statistical downscaling method. Air temperature fields are spatially downscaled from 2.0° to 1.0° using a linear interpolation and temporally downscaled from 6-hourly to 3-hourly using a bilinear interpolation. For the 1-month and 2-month forecasts, we use a station observation-based monthly dataset at 0.5° latitude/longitude resolution recently developed at the Climate Prediction Center, National Centers for Environmental Prediction [Fan and Van Den Dool, 2008]. This dataset combines station observations collected from the Global Historical Climate Network and the Climate Anomaly Monitoring System (GHCN + CAMS) and uses an interpolation technique accounting for topography effects.

4.4 Results

4.4.1 Forecast Skill

4.4.1.1 Soil Moisture

Before evaluating the 2-m air temperature and precipitation forecast skill, we verify whether the PAR initialization improves the soil moisture forecasts. Recall that the only available
observational dataset of soil moisture for the study period (1986-1995) are in-situ measurements over Illinois. Figure 4.2 displays the vertical profile of average soil moisture anomalies over Illinois for the observations (top panel), the PAR forecasts (middle panel) and the control forecasts (bottom panel). The anomalies are calculated with respect to the monthly climatology based on the study period. The forecasts results shown here are at 1-month time scale.

Figure 4.2 shows a few important discrepancies between the two series of forecasts. For instance, the PAR forecasts capture very well the amplitude and the duration of both the 1986 drought (dry anomalies in red) and the 1993 flood (wet anomalies in blue) that occurred in Illinois, while the control forecasts underestimate both. On the left hand-side of Figure 4.2 the temporal correlation square ($r^2$) of soil moisture anomalies between each series of forecasts and the observations are shown for each soil layer. It is clear that the values of $r^2$ for the PAR forecasts are higher than those for the control forecasts throughout the soil column. One can notice that the difference of $r^2$ between the two series of forecasts increases with soil depth; in the top layer, the value of $r^2$ goes from 0.15 (control forecasts) to 0.19 (PAR forecasts) corresponding to an increase of only of 4% of the fraction of explained variance; in the deepest layer (200 cm), the value of $r^2$ goes from 0.05 (control forecasts) to 0.4 (PAR forecasts) corresponding to a large increase of 35% of the fraction of explained variance. Another very interesting difference between the two series of forecasts is noticeable at soil depths deeper than 100 cm. The soil moisture anomalies of the control forecasts fade at about 100 cm while these of the PAR forecasts extend down to 200 cm. This indicates that, below 100 cm soil depth, the PAR forecasts ”remember” the soil moisture anomalies of the soil moisture analysis and thus better match with the observations than control forecasts. This explains why the skill of the assimilated forecasts increases with depth.

To conclude, we find that the soil moisture analysis improves significantly the 1-month soil moisture forecasts. In addition, the comparison of the two series of forecasts at deep layers (lower than 100 cm) emphasizes the memory effect of soil moisture. This result is therefore promising for enhancing the near surface air temperature and precipitation forecast skill.
Figure 4.2: Vertical profile of average soil moisture anomalies over Illinois for the observations, the PAR forecasts (real. for) and the control forecasts (ctl for). 7 layers are shown: top-10, 10-30, 30-50, 50-70, 70-90, 90-110 and 110-200 cm. The left hand side represents the anomaly correlation between each set of forecasts and the observations. The gray area indicates that the anomaly is not statistically significant at the 99% confidence interval.

4.4.1.2 Near Surface Air Temperature (T2m)

Figure 4.3 shows the spatial distribution of the T2m anomaly correlation (r) between each series of forecasts and the observations. The anomalies are calculated with respect to the study period. Three different time-scales are analyzed from short-term (15-day time scale) to subseasonal (2-month time scale). Correlation coefficients between -0.25 and 0.25 are not
statistically significant at the 95% confidence interval. The confidence intervals are calculated using *Emery and Thomson* [1998]. The right side of Figure 4.3 represents the difference of anomaly correlation squared ($r^2$) between the PAR and control forecasts. As mentioned earlier, the difference measures the increase (or decrease) of the fraction of explained variance attributed to the realistic soil moisture initial conditions. The gray thick contours indicate where the correlation differences are statistically significant at the 95% confidence interval.

It is clear that the maps of $r^2$ difference show a positive impact on the 2-m air temperature forecasts across most of the continental U.S. for all time scales (Figure 4.3). Indeed, at all time scales, the values of $r^2$ difference go up to $+0.24$. This indicates that the soil moisture analysis contributes to 24% of the fraction of explained variance. One can notice that, at the 15-day and 1-month time scales, the positive values of $r^2$ difference are located mostly over the Rocky Mountains and the Great Plains. This is consistent with the study conducted by GLACE-1 showing a strong land-atmosphere coupling strength in these regions (Figure 4.1). At the 2-month time scale, a northward shift of the forecast skill maxima is shown. This is most likely due to the soil moisture memory effect; it is well known that the time of soil moisture memory increases with latitude [*Delworth and Manabe 1988; Koster and Suarez 2001; Schlosser and Milly 2002; Wu and Dickinson 2004, Seneviratne et al. 2006*].

To better understand the causes of the T2m forecast skill increase, Figure 4.4 compares the timeseries of the forecasted T2m anomalies with the forecasted soil moisture anomalies averaged over the box centered in grid cell 5 (indicated in Figure 4.6a). It is shown that the T2m anomalies of the PAR forecasts diverge the most from those of the control forecasts from August 1991 until August 1992, and become closer to observations. During this same time period, we can notice that the vertical profile of soil moisture anomalies shows a shift from dry anomalies (positive values in red) in the control forecasts to strong wet anomalies (negative values in blue) in the PAR forecasts. This strongly suggests that it is the wet soil moisture anomalies that are responsible for greatly reducing the T2m anomalies (from positive to negative values) and thus are responsible for pushing those anomalies close to observations.
**Figure 4.3:** 2-m air temperature anomaly correlations against observations of the PAR forecasts (real. for, left column) and the control forecasts (ctl for, middle column) for 15 days, 1-month and 2-month forecasts. The right column displays the percentage of skill increase (or decrease) in terms of fraction of explained variance attributed to the realistic soil moisture initialization.

### 4.4.1.3 Precipitation

Figure 4.5 is the same as Figure 4.3 but for precipitation forecasts. The statistical tools used here are the same as used in Figure 4.3. It is clear that the results are not as optimistic as those of the T2m forecasts. At all time scales, there is no statistically significant effect of the soil moisture analysis on precipitation forecasts. This result can be explained by two factors: 1) it is well known that to forecast accurately precipitation during the boreal summer is very
Figure 4.4: Time series of the average forecasted anomalies of T2m (top panel) and soil moisture vertical profile of PAR forecasts (real. for, middle panel) and the control forecasts (ctl for, bottom panel) over the box centered in grid cell 5 (Figure 4.6a). The time scale of the forecasts considered here is the average over the first month.

challenging [Olson et al. 1995; Chien et al. 2005]; 2) through the modification of sensible fluxes, the impact of soil moisture fields on the T2m state is rather direct while several intermediate physical processes occur before the latent fluxes can affect the generation of precipitation. Therefore, a weaker impact of the soil moisture analysis on the precipitation forecasts is not surprising. It also is plausible that the increase in precipitation forecast skill is limited by the two following model-dependent characteristics: 1) the too coarse spatial resolution used in this study (T63 ∼ 200 km) can restrain the climate model to capture any soil moisture-convection feedback; or 2) the precipitation response to soil moisture conditions in the FSU/COAPS climate model could be too weak. This last factor is investigated in the
4.4.2 Model Response

Using the 60 forecasts, here we concentrate on the response of T2m and precipitation to a change in soil moisture level. We saw earlier that the soil moisture impact on the T2m (via the modification of surface fluxes) is rather direct and its impact on precipitation is rather indirect since several intermediate physical processes occur before the latent fluxes (modified to soil moisture) can affect the generation of precipitation. Therefore, a local impact on
T2m and a regional impact on precipitation are assumed. Then, we analyze the response to a change of soil moisture level at six different grid points for T2m shown in Figure 4.6a and over six average areas for precipitation (in the black boxes shown in Figure 4.6b). The yellow zones indicate when the model response to a change of soil moisture level is physically consistent (i.e., positive response). In other words, the yellow zones indicate when an increase (decrease) of soil moisture level leads to a decrease (increase) of T2m. For precipitation, the yellow zones indicate when an increase (decrease) of soil moisture level leads to an increase (decrease) of precipitation via a feedback mechanism. The ratio located on the top right of each panel represents the number of forecasts with a positive response out of the total number of forecasts. The change $\delta(x)$ is defined as follows:

$$
\delta(x) = \frac{x(\text{PAR forecasts}) - x(\text{control forecasts})}{x(\text{PAR forecasts})}
$$

$x$ being one of the following variables: soil moisture, precipitation or air temperature.

**Figure 4.6:** Locations of the grid cells (black crosses) and areas (black boxes)Overlaying the difference of the time anomaly correlations square between the PAR and control forecasts at 1-month time scale for a) air temperature and b) precipitation forecasts displayed in Figures 4.3 and 4.5 respectively.

Figure 4.7 displays the relationship between a change (defined in equation 4.3) of soil moisture level and a change of T2m at 0 time lag. This Figure reveals a linear relationship between these two variables at all considered grid cells. In addition, most of the 1-month T2m forecasts sustain a positive response to a soil moisture change meaning a soil moisture
increase (decrease) induces a 2-m air-temperature decrease (increase). The magnitude of the response is the largest at grid cells 2, 3, 4 and 5 where the sensitivity of the T2m (values of $delta(T2m)$) is about twice larger than at the other grid cells and where most of the forecasts are located in the yellow zones (high ratios). One can notice that those grid cells (2, 3, 4 and 5) are located in regions where the increase in 2-m air temperature forecast skill is the largest (Figure 4.6a). Thus, this suggests that the regions of maximum skill increase are located where the model response of T2m to a soil moisture change is the strongest.

It is evident that precipitation affects soil moisture conditions. The question of whether soil moisture conditions affect precipitation via a feedback mechanism is less obvious. To answer this question, Figure 4.8 displays the precipitation response to a soil moisture change of the earlier month (-1 month lag). First, it is interesting to notice that, unlike the T2m response, the precipitation response does not appear to be linear. The comparison of the six different regions reveals that the precipitation response depends on the region. Region 1 hardly shows any positive response (ratio = 10/60). Regions 3, 5 and 6 show a positive response for about half of the forecasts. Region 2 sustains a positive response for most of the forecasts (ratio = 53/60) but the amplitude of the response is weak. Region 4 is particularly interesting because: 1) it has a high number of forecasts showing a positive response (ratio = 39/60); and 2) the precipitation sensitivity to a soil moisture change is large (values of $delta(P)$ up to -200%). In addition, one can notice that the precipitation sensitivity is larger to a dry soil moisture change than a wet soil moisture change. This could suggest an interesting result that the soil moisture-convection feedback is stronger over a dry soil than a wet soil and therefore could play a greater role in maintaining a drought than a flood. We will further investigate this result in section 4.4.4.
Figure 4.7: Relationship between a change of the forecasted 2-m air temperatures and a change of the forecasted soil moisture (sm). Delta(sm) and delta(T2m) are defined in equation 4.3.

Figure 4.8: Relationship between a change of the forecasted precipitation (P) and a change of the forecasted soil moisture (sm). Delta(sm) and delta(P) are defined in equation 4.3.
4.4.3 Results of GLACE-2

As mentioned in section 4.1, this study and the ongoing GLACE-2 model intercomparison experiment have similar goals. Using several climate models, including the FSU/COAPS model, GLACE-2 also aims at investigating the contribution of a realistic soil moisture initialization to the subseasonal forecasting skill of precipitation and 2-m air temperature. In order to compare the forecast skill results of the land-atmosphere FSU/COAPS model with other land-atmosphere models, here we present some unpublished results produced by the GLACE-2 team. This comparison allows us to verify whether the results of the land-atmosphere FSU/COAPS model are comparable with other models. Only five climate models are displayed here: COLA, NCEP, ECMWF, KNMI and FSU/COAPS. The Multi-Model Analysis (MMA) represents the mean of the five models. Figures 4.9 to 4.12 show the values of T2m $r^2$ for the control forecasts (series-2), the realistically initialized forecasts (series-1) and their difference at 1-15, 15-30, 31-45 and 45-60 days lead time respectively. Figures 4.13 to 4.16 are the same Figures as 4.9 to 4.12 but for precipitation forecasts. To make these Figures, 100 forecasts are used (instead of the 60 used in this study) for the calculation of $r^2$; the forecast starting dates are the same as those used in this study (described in section 4.2) but include the 1st and 15th of the months of April and May in each of the ten years spanning 1986 to 1995.

For T2m, all land-atmosphere models show an increase in forecast skill due to the realistic soil moisture initialization (i.e. soil moisture analysis) at most lead times. However, a large diversity in the locations of maxima among models is seen. For precipitation, the pessimistic results of the land-atmosphere FSU/COAPS model that we earlier found in Figure 4.5 are consistent with the other climate models at all lead times. There is no significant impact of the realistic soil moisture initialization (i.e. soil moisture analysis) for all models. This brief model intercomparison indicates that the results of the FSU/COAPS climate model are comparable with these of the other models participating in GLACE-2 for both T2m and precipitation forecasts.
Figure 4.9: 2-m air temperature anomaly correlation square ($r^2$) against observations of the series-2 (left column) and the series-1 (middle column) for 1-15 days time scale. The right column displays the difference between the series-1 and series-2. The series-1 is equivalent to our PAR forecasts and the series-2 is equivalent to our control forecasts.
Figure 4.10: Same Figure as 4.9 but for 16-30 time scale
Figure 4.11: Same Figure as 4.9 but for 31-45 time scale.
Figure 4.12: Same Figure as 4.9 but for 46-60 days time scale.
Figure 4.13: Precipitation anomaly correlation square (r²) against observations of the series-2 (left column) and the series-1 (right column) for 1-15 days time scale. The right panels display the difference between the series-1 and series-2. The series-1 is equivalent to our PAR forecasts and the series-2 is equivalent to our control forecasts.
Rainfall Skill for 1986–1995 (Higgins day 16–30)

Series 2  |  Series 1  |  Difference

**COLA**

**NCEP**

**ECMWF**

**KNMI**

**FSU**

**MMA**

**Figure 4.14:** Same Figure as 4.13 but for 16-30 days time scale.
Figure 4.15: Same Figure as 4.14 but for 31-45 days time scale.
Figure 4.16: Same Figure as 4.15 but for 46-60 days time scale.
4.4.4 Extreme Events

Several studies have suggested that SSTs trigger extreme events while soil moisture is responsible for maintaining and/or intensifying them [Trenberth and Branstator, 1992; Atlas et al., 1993; Bosilovich and Sun, 1999; Hong and Kalnay, 2002, etc.]. During the 10-year study period (1986-1995), two major extreme events occurred over the central U.S. in the boreal summer: the 1988 drought and the 1993 flood. The 1988 U.S. drought was the worst since the Dust Bowl (1930 to 1936), covering 40% of the country at the peak of the drought in July as measured by the Palmer Drought Severity Index. The Great 1993 U.S. Flood was unusual as far as the magnitude, the large impacted area, and the duration (May to September) are concerned. Those two hydrological extreme events had profound negative impacts on economic, environmental, and social sectors. The damages caused by each event cost tens of billions of dollars with 5,000 to 10,000 lives lost in the 1988 drought and over 50 lives lost in the 1993 flood. The fact that these two natural disasters occurred during a strong La Niña event (1988) and a rather unusual El Niño event (1993) suggests that the ENSO events play an important role in the onset of these types of extreme events [Trenberth and Branstator 1992; Trenberth and Guillemot 1996]. Indeed, the anomalous SSTs that prevail in the tropical Pacific Ocean during ENSO events are responsible for major shifts in the position of the jet stream, which in turn affects the storm track and thus the climate over the U.S. An El Niño event tends to shift the jet stream southwards, increasing precipitation over the continental U.S. Conversely, a La Niña event is associated with large upper-level anticyclonic height anomalies over the continental U.S., which push the jet stream further north into central Canada and reduce precipitation over the continental U.S. The role of ENSO in triggering large-scale atmospheric circulation anomalies favorable for the formation of drought and flood over the U.S. is well recognized [Trenberth and Branstator, 1992; Trenberth and Guillemot, 1996; Atlas et al., 1993; Bosilovich and Sun, 1999; Hong and Kalnay, 2002, etc.]. However, it is believed that soil moisture conditions are also crucial for maintaining and/or intensifying these extreme events [Atlas et al., 1993; Bosilovich and Sun, 1999; Hong and Kalnay, 2002]. In this section, we investigate the impact of the soil moisture analysis produced in Chapter 3 on the forecasts of these two natural disasters. The metric to evaluate these particular forecasts is the square of the spatial correlation ($r^2$) between each series of forecasts and observations and the results are displayed in Table 4.3.

66
To quantify the improvement of skill for those particular years, the difference of $r^2$ between the PAR and control forecasts is computed. The difference of $r^2$ is set to zero if one of the correlations is negative or not statistically significant.

Figure 4.17 presents the 1-month forecasts of the average precipitation, T2m and total soil moisture (1.5 m depth) anomalies for the PAR forecasts, control forecasts and observations over the boreal summer months (JJA) of the year 1988. For the map of precipitation observations, the white areas in the northern part of the U.S. denote undefined data. Figure 4.19a shows the difference of each variable between the PAR and control forecasts of the year 1988. The comparison of precipitation anomalies shows some differences between the PAR and control forecasts over the Midwest (Figure 4.17). The PAR forecasts eliminate the wet precipitation anomalies over the Great Plains and increase the amplitude of dry precipitation anomalies below the Great Lakes by more than 1.4 mm.dy$^{-1}$. Therefore, the PAR forecasts better compare with the observations than the control forecasts as far as intensity and spatial distribution. Table 4.3 shows an increase in skill of 11% of the fraction of explained variance for the month of June. Recall that no significant impact was found when considering all years between 1986 and 1995 to calculate the forecast skill (Figure 4.5). The statistically significant skill increase of 11% suggests that the soil moisture analysis is able to improve the precipitation forecasts during extreme dry events such as the drought year of 1988. For T2m forecasts, the PAR forecasts extend the dry anomalies further west and intensify them of about 1K.d$^{-1}$. Table 4.3 shows a great increase in the fraction of explained variance (14% in June and 13% in July) for the T2m forecasts. Due to the lack of global observations of soil moisture, the soil moisture anomalies of both series of forecasts are compared to the GSWP-2 (described in Chapter 3). Recall that GSWP-2 is an offline multi-model analysis of land surface models driven by atmospheric observation-based forcing and is used here as a proxy for global soil moisture observations. Across most of the U.S., the PAR forecasts better reproduce the anomaly pattern and intensity of GSWP-2 than the control forecasts. With respect to GSWP-2, the values of $r^2$ difference show a significant increase in the months of June (14%) and July (19%).
Figure 4.17: Average precipitation (top panels), the 2-m air temperature (middle panels) and the total soil moisture anomalies (bottom panels) of the PAR forecasts (real. for, left column), the control forecasts (ctl for, middle column) and the observations (right column) for the summer months (JJA) of the year 1988. The precipitation anomalies are in mm.d\(^{-1}\), the 2-m air temperature are in K.d\(^{-1}\) and the volumetric soil moisture m\(^3\).m\(^{-3}\).month\(^{-1}\).

Figure 4.18 is the same as Figure 4.17 but for the flood year of 1993. The Table 4.4 uses the same statistical approach as in the Table 4.3. The comparison of precipitation anomalies shows that the wet anomalies are very localized under the great lakes in the control forecasts while they extend further to the Midwest in the PAR forecasts as seen in the observations. The difference of \(r^2\) values shows a very small increase (2\%) of the fraction of explained variance in June (Table 4.4). Although both series of forecasts underestimate the
Table 4.3: Spatial correlations squared ($r^2$) over the U.S. between each series of forecasts and the observations for each boreal summer month of the drought year of 1988. The difference of $r^2$ between the PAR forecasts and the control forecasts was also computed. The values of correlation or the difference of correlation are in bold when they are statistically significant at the 95% confidence interval.

<table>
<thead>
<tr>
<th></th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAR for</td>
<td>ctl for</td>
<td>diff</td>
</tr>
<tr>
<td>Precipitation</td>
<td>0.14</td>
<td>0.03</td>
<td>0.11</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.15</td>
<td>0.017</td>
<td>0.14</td>
</tr>
<tr>
<td>Soil moisture</td>
<td>0.18</td>
<td>0.03</td>
<td>0.14</td>
</tr>
</tbody>
</table>

amplitude of the overall wet anomalies in the Midwest, a small increase of its precipitation intensity by up to 0.8 mm.d$^{-1}$ is noticeable in the PAR forecasts over the control forecasts (Figure 4.19b). The comparison of T2m anomalies clearly shows an improvement in terms of intensity (Figure 4.19b); the amplitude of the anomalies is increased in the PAR forecasts by 1.4 K.d$^{-1}$ and matches better with that of the observations. The fraction of explained variance attributed to the soil moisture analysis is significantly increased by 6% in July and 7% in August (Table 4.4). Finally, the comparison of soil moisture anomalies shows many differences between both series of forecasts. The control forecasts do not pick up the wet anomalies over the Midwest and exhibit a dry anomaly above the Great Lakes that is not observed in GSWP-2. Despite a generally weaker intensity of the wet anomalies compared to GSWP-2, the spatial distribution of the PAR forecasts matches better with that of GSWP-2. The values of $r^2$ difference show a significant increase of the fraction of explained variance for each boreal summer month with a maximum in August of 32% (Table 4.4).
Table 4.4: Same as Table 4.3 but the flood year of 1993.

<table>
<thead>
<tr>
<th></th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAR for</td>
<td>ctrl for</td>
<td>diff</td>
</tr>
<tr>
<td>Precipitation</td>
<td>0.06</td>
<td>0.04</td>
<td><strong>0.02</strong></td>
</tr>
<tr>
<td>Temperature</td>
<td>0.48</td>
<td>0.46</td>
<td><strong>0.02</strong></td>
</tr>
<tr>
<td>Soil moisture</td>
<td>0.10</td>
<td>0.01</td>
<td><strong>0.09</strong></td>
</tr>
</tbody>
</table>

Figure 4.18: Same as Figure 4.17 but for the flood year of 1993.
Figure 4.19: Anomalies difference between the PAR forecasts and the control forecasts of precipitation (top), 2-m air temperature (middle) and soil moisture (bottom) for a) the 1988 drought (left column) and b) the 1993 flood (right column).

When we compare the results of the two extreme event forecasts, the 1988 drought forecast shows overall greater improvements than the 1993 flood forecast. This suggests that soil moisture contributes more to the development of a drought than that of a flood. This
result is in agreement with the study by Weaver et al., [2009]. To further investigate this result, we assess the evaporation sources of precipitable water following the methodology by Dirmeyer and Brubaker [1999]. The evaporation sources are calculated using back trajectories using the fully implicit technique of Merrill et al. [1986]:

\[
x^{n-1} = x^n + \frac{\tau}{2}[u^n + u^{n-1}]
\]

\[
y^{n-1} = y^n + \frac{\tau}{2}[v^n + v^{n-1}]
\]

where \(x\) and \(y\) are the locations of an air parcel from evaporation, \(u\) and \(v\) are the zonal and meridional components of the wind, \(n\) denotes the time step and \(\tau\) (1 day) the time interval. The air parcel is traced 15 days back in time only where precipitation occurred in the target region. Here, the target region the central U.S. because it is where the 1988 drought and the 1993 flood occurred. The fraction of water \(S(x,y)\) falling on the target region \(A\) that originates from a grid cell \((x,y)\) is calculated as follows:

\[
S(x,y) = \frac{E(x,y)}{P_A}
\]

where \(E(x,y)\) is the water mass contribution of evaporation from grid cell \((x,y)\) and \(P_A\) is the total precipitation over the region \(A\). When the grid cell \((x,y)\) is inside the region \(A\), \(S(x,y)\) represents the recycling ratio. The recycling ratio is defined as how much local evaporation contributes to precipitation. When this ratio is high, it is a good indicator of soil moisture-precipitation feedback. More details are given in Dirmeyer and Brubaker [1999].

Figure 4.20 shows the evaporative sources \(S(x,y)\) falling over the central U.S. (black rectangle) during the summer months (June, July and August). The comparison of these two years clearly shows that the evaporative sources for the year 1988 are mostly terrestrial while for the year 1993 the evaporative sources are mostly oceanic from the Gulf of Mexico and the Caribbean Sea. The recycling ratio (values of \(S(x,y)\) in the black rectangle of Figure 4.20) is greater in the 1988 drought than the 1993 flood suggesting that dry anomalies are further maintained by regional soil moisture-precipitation feedback mechanism than wet anomalies. Therefore, stronger feedbacks occurring during an extreme dry year than an extreme wet year could explain why the soil moisture analysis has more impacts on the drought year of 1988 than a flood year of 1993 found in Figure 4.17 and 4.18. This result is also in very
good agreement with that of *Dirmeyer and Brubaker*, [1999]. Unlike this study, *Dirmeyer and Brubaker*, [1999] calculated it with respect to the rain falling over different mid-latitude regions and using hourly data.

**Figure 4.20:** Evaporative sources (kg m \(^{-2}\)) for rain falling over the target region (black rectangle) during June-August of the year a) 1988 and b) 1993. The difference of evaporation sources between the year 1988 and 1993 is displayed in panel c).
4.5 Summary and Discussion

In this chapter, using the land-atmosphere FSU/COAPS model, we investigate the contribution of a realistic soil moisture initialization as close to the truth as possible (using the soil moisture analysis produced in Chapter 3) on summertime 2-m air temperature and precipitation forecasts over the continental U.S. between 1986 and 1995. We find that the 2-m air temperature forecast skill is significantly increased across most of the U.S. (up to a 24% increase in the fraction of explained variance) at a short-term to subseasonal time scale. At 15-day and 1-month time scales, the regions where the maximum skill increase for the 2-m air temperature forecasts is consistent with the "hot-spots" of GLACE-1, defined as the regions with strong land-atmosphere coupling strength. At a 2-month time scale, the maximum increase is shifted northward, which emphasizes the effect of soil moisture memory, whose time length increases with latitude. In the other hand, no skill increase is found for the precipitation forecasts at any time-scale. The GLACE-2 model intercomparison experiment shows that the forecast skill results of the land-atmosphere FSU/COAPS model is comparable with the other participating climate models of GLACE-2.

Because it is thought that soil moisture plays an important role in the development of extreme hydrological events, we also focus on the particular summer drought year of 1988 and flood year of 1993 in this section. The results show an overall improvement in intensity and spatial coverage of both extreme event forecasts. In particular, unlike the result found earlier of an absence of precipitation forecast skill increase between 1986 and 1995, the year 1988 sees a statistically significant improvement in intensity and spatial distribution of precipitation forecasts. This may suggest that the soil moisture analysis is capable of increasing precipitation forecast skill only during dry years. This result is consistent with a previous study by Weaver et al., [2009]. In general, a greater skill increase is found for the 1988 drought forecasts than for the 1993 flood forecasts. To better understand and compare the influence of soil moisture on precipitation generation between a dry and wet year, we analyze the evaporative sources of these two extreme events using the back-trajectory methodology of Dirmeyer and Brubaker, [1999]. The result indicates that the evaporative sources of the 1988 drought are regional while these of the 1993 flood are more remote coming from the Gulf of Mexico and the Caribbean Sea. This agrees very well
with the results found by *Dirmeyer and Brubaker* [1999] which conducted a similar analysis but using NCEP reanalysis data. This evaporative sources analysis indicates that the soil moisture-precipitation feedback mechanism plays a greater role during a dry year than a wet year. It thus explains why the soil moisture analysis favors the improvement of the 1988 drought forecast over that of a flood forecast. In addition, this result is consistent with a larger precipitation sensitivity to a dry soil than a wet soil found in Figure 4.8.
CHAPTER 5

CONCLUSION

This dissertation presents the development and validation of a soil moisture analysis and its impact on the short-term to subseasonal forecasting skill of 2-m air temperature and precipitation. This soil moisture analysis allows to initialize the climate forecasts to values that are as close to the truth (i.e. realistic) as possible. Because of its long memory and its role in controlling the surface water and energy budget, soil moisture is today widely recognized as having the potential to improve summer forecasting skill. However, because of a lack of global observations of soil moisture, there has been no scientific consensus on the contribution of realistic soil moisture initial conditions to climate forecasting skill. In this study, the soil moisture initial conditions (i.e. analysis) are generated using a physically consistent Precipitation Assimilation Reanalysis (PAR) technique. This technique consists mainly of nudging precipitation in the atmosphere component of a land-atmosphere model in a coupled mode by adjusting the vertical air humidity profile based on the difference between the rate of the model-derived value and the observed precipitation value.

We find that the implementation of the PAR technique into the coupled land-atmosphere FSU/COAPS model produces model-derived soil moisture estimates that are in good agreement with observations over Illinois deep into the soil as far as the seasonal cycle and the monthly anomaly variability (anomaly correlations > 0.5 at each grid point in Illinois). The comparison on a global-scale with a proxy for soil moisture observations (i.e. the GSWP-2 multi-model analysis) in the top 1.5 m of soil also gives optimistic results. During the boreal summer, the spatial distribution, absolute amount and anomaly variability of soil moisture (anomaly correlations up to 0.8 across most of the globe) are much closer to GSWP-2 compared to a free run (i.e. without assimilation). Not only the soil moisture estimates
but also other land surface variables, such as soil evaporation and surface runoff estimates are close to the values of GSWP-2. However, one should be cautious in interpreting those results. Because the GSWP-2 is a multi-model soil moisture analysis product, it is not necessarily always accurate, especially if there is a high variability among models, which is the case for the GSWP-2 surface runoff estimate. Nevertheless, GSWP-2 offers the best proxy for land surface observations because it merges state-of-the-art LSMs driven by an atmospheric observation-based forcing. Since the NCEP-R2 reanalysis uses a similar soil moisture land assimilation system as that used in this study, we also compare our results with this reanalysis. The soil moisture analysis in the NCEP-R2 reanalysis is produced by assimilating observed precipitation into a land-atmosphere model in a coupled mode (similar to this study). However, because precipitation is assimilated at the land surface (unlike in this study, not in the atmosphere component of the model), the resulting soil moisture analysis is not physically consistent with the surface and radiative fluxes when the atmospheric component of the reanalysis model presents an error. The comparison shows that the soil moisture analysis produced in this study 1) better correlates with observations than that of the NCEP-R2 over Illinois and 2) is most of the time closer to GSWP-2. This finding suggests that the use of a physically consistent land assimilation system is important in producing soil moisture initial conditions that are as close to truth as possible. Therefore, in this study, we have developed a new global soil moisture analysis dataset that is comparable with a benchmark in global soil moisture analysis (i.e. GSWP-2) and thus can be used for many land surface studies (crop modeling, water resource management, soil erosion, climate variability, etc.).

Many numerical sensitivity studies have shown that soil moisture can potentially play an important role in controlling the variability of summertime near surface variables in the mid-latitudes. Therefore, in this study, we analyze the contribution of the above described soil moisture analysis on 2-m air temperature and precipitation forecasts over the continental U.S. during the boreal summer months (June, July and August) between 1986 and 1995. We find that the 2-m air temperature forecast skill is significantly increased (up to 24% of the fraction of explained variance) across most of the U.S. at a short-term (2 weeks) to subseasonal (2 months) time scale. At 15-day and 1-month time scales, the regions of maximum skill increase for the 2-m air temperature forecasts are consistent with the "hot-spots" found in the
GLACE-1 experiment, which are defined as the regions with strong land-atmosphere coupling strength. At the 2-month time scale, the maximum skill increase is shifted northward, which emphasizes the effect of soil moisture memory, whose time length increases with latitude. On the other hand, no skill increase is found for the precipitation forecasts at all study time-scales. This lack of skill for precipitation forecasts can be explained by the following factors: 1) it is well known that to forecast accurately summertime precipitation is a very challenging task; 2) through the modification of sensible fluxes, the impact of soil moisture on the 2-m air temperature state is rather direct while several intermediate physical processes occur before the latent fluxes can affect the generation of precipitation; 3) in these results all years between 1986 and 1995 are considered in the calculation of forecast skill; it is possible that soil moisture can affect precipitation generation significantly only in particular years associated with extreme hydrological events, such as drought and flood.

Since several sensitivity studies have demonstrated that soil moisture could potentially play a role of intensifying and/or prolonging extreme hydrological events [Trenberth and Branstator, 1992; Atlas et al., 1993; Bosilovich and Sun, 1999; Hong and Kalnay, 2002, etc.], we also analyze the impact of the soil moisture analysis on the summertime forecasts of two particular years of 1988 and 1993. During the summer of those two particular years, a drought and flood respectively caused exceptional damages over the central U.S. For the precipitation forecasts, the results show a statistically significant skill increase (11% of the fraction of explained variance) for the year of 1988 but no skill increase for the year of 1993. For the 2-m air temperature and soil moisture forecasts, the comparison of the two extreme hydrological event forecasts shows greater improvements for the year of 1988 than for the year of 1993. For instance, the fraction of explained variance for 2-m air temperature forecasts is increased by 14% during the 1988 drought compared to only 6% during the 1993 flood. This suggests that soil moisture contributes more to the development of a drought than a flood. These results are consistent with the studies by Dirmeyer and Brubaker, [1999] and Weaver et al., [2009]. To better understand and compare the role of soil moisture on the land-atmosphere feedback mechanism between a dry (i.e. year of 1988) and wet year (i.e year of 1993), we analyze the evaporative sources of these two extreme events using the back-trajectory methodology of Dirmeyer and Brubaker, [1999]. The results indicate that the evaporative sources of the 1988 drought are more local while these of the 1993 flood are
more remote coming from the Gulf of Mexico and the Caribbean Sea. This suggests that
the soil moisture-precipitation feedback mechanism plays a greater role during an extremely
dry year than an extremely wet year and thus explains why the soil moisture analysis favors
the improvement of the 1988 drought forecast over that of a flood forecast. This result
is in very good agreement with the results found by Dirmeyer and Brubaker [1999], who
conducted a similar analysis using NCEP reanalysis data. In addition, it is consistent with
larger precipitation sensitivity to a dry soil than a wet soil found in this study (Figure 4.8).

Finally, as part of my Ph.D research, using the coupled land-atmosphere FSU/COAPS
model, I have joined the GLACE-2 international model intercomparison project. The
participation of the land-atmosphere FSU/COAPS model in GLACE-2 presents the unique
opportunity to compare our model results with those of the other participating land-
atmosphere models. Unlike almost all the participants in GLACE-2 that use an offline land
assimilation system, we produce a soil moisture analysis using a land-atmosphere model in a
coupled mode (i.e. PAR technique). A brief comparison with GLACE-2 results shows that
the FSU/COAPS model produces comparable results of forecasting skill of precipitation and
2-m air temperature with the other participating models. Therefore, this suggests that the
overall forecast skill increase attributed to a soil moisture initialization that is as close to
the truth as possible does not depend on the land assimilation technique or on the climate
model.
APPENDIX A

SOIL MOISTURE MODEL

This Appendix aims to describe in greater detail the computation of the soil moisture state in the Community Land Model (CLM2). The Figure A.1 displays all the variables included in the hydrological cycle over land.

![CLM2 Soil Moisture Model](image)

**Figure A.1:** CLM2 Soil Moisture Model.
A.0.1 soil moisture at surface

At the surface, the soil moisture is calculated based on the surface water balance:

\[ q_{soilm} = q_{precip} + q_{drip} + q_{msnow} - q_{runf} - q_{gevap} \]  \hspace{1cm} (A.1)

where \( q_{precip} \) is the precipitation reaching the surface, which either infiltrates the soil or evaporates \( q_{gevap} \) or runs off the land surface \( q_{runf} \). Other minor sources of soil moisture are the canopy drip \( q_{drip} \) and the melting snow \( q_{msnow} \) during the spring season. The PAR technique directly modifies \( q_{precip} \) by its precipitation nudging and \( q_{gevap} \) by its dynamical nudging.

- **runoff**

The runoff in the CLM2 model is a hybrid of the runoff scheme of BATS (Dickinson et al. 1993) and TOPMODEL (Beven and Kirkby 1979). It is calculated based on the total amount of water reaching the soil \( q_{wat.grd} \) (sum of precipitation, canopy drip and melting snow), and the fractional saturated area \( F_{sat} \), that is directly related to the level of soil moisture:

\[ q_{runf} = F_{sat} - q_{wat.grd} + (1 - F_{sat})w_s q_{wat.grd} \]  \hspace{1cm} (A.2)

Where, \( w_s \) is the layer depth weighted soil moisture over the first 3 layers (91 cm). Since \( q_{runf} \) depends on soil moisture content, it is indirectly affected by the PAR technique.

- **ground evaporation**

The ground evaporation \( q_{evap} \) is related to the gradient of specific humidity between the ground and the atmosphere as follows:

\[ q_{gevap} = \frac{\rho(q_a - q_g)}{r_{aw}} \]  \hspace{1cm} (A.3)

Where \( r_{aw} \) is the aerodynamical resistance of a vegetated soil that depends on the strength of the wind, and \( q_a \) and \( q_g \) are the atmospheric and ground specific humidity, respectively. While the \( r_{aw} \) and \( q_a \) are directly modified by the PAR technique, the \( q_g \) is modified through the soil moisture amount.
A.0.2 soil moisture at deeper layer

At deeper layers \( (z>1) \), the soil moisture is governed by the difference of vertical moisture between the upper layer \( (q_{z+1}) \) and lower layer \( (q_z) \) and the transpiration \( (q_{\text{transp},z}) \). The index \( j \) indicates the type of vegetation (or PFT).

\[
q_{\text{soilm},z} = \frac{\Delta t}{\Delta z} (q_{z+1} - q_z) - q_{\text{transp},z} \tag{A.4}
\]

The vertical moisture flux \( q_z \), is described by the following Darcy’s law:

\[
q_z = -k \frac{\delta q_{\text{soilm},z}}{\delta z} \tag{A.5}
\]

where \( k \) is the hydraulic conductivity.
REFERENCES


Marie Boisserie

In 2000, Marie Boisserie completed a Bachelor’s degree in Mathematics at the Paris Diderot University, France and in 2001 she completed a Bachelor’s degree in Physics at the ”Pierre et Marie Curie” University, Paris, France. She obtained her Master’s degree in summer of 2003 in Oceanography, Meteorology and Environment (OME), also at the ”Pierre et Marie Curie”, Paris. She enrolled in the doctoral program at FSU in the fall of 2005. Marie’s research interests include land-atmosphere interactions, soil moisture measurement, modeling, and data assimilation.