ARISTOTLE UNIVERSITY OF THESSALONIKI
FACULTY OF SCIENCES
SCHOOL OF GEOLOGY

POST-GRADUATE PROGRAMME OF METEOROLOGY, CLIMATOLOGY AND ATMOSPHERIC ENVIRONMENT

EFTHYMIOS SERPETZOGLOU
B.SC. PHYSICS
M.SC. METEOROLOGY AND PHYSICAL OCEANOGRAPHY

DEVELOPMENT AND APPLICATION OF ASSIMILATION TECHNIQUES OF HYDROMETEOROLOGICAL REMOTELY-SENSED DATA IN METEOROLOGICAL AND LAND-SURFACE MODELS

A DOCTORATE DISSERTATION

THESSALONIKI 2018

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΕΩΡΟΛΟΓΙΑΣ, ΚΛΙΜΑΤΟΛΟΓΙΑΣ ΚΑΙ ΑΤΜΟΣΦΑΙΡΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

ΕΥΘΥΜΙΟΣ ΣΕΡΠΕΤΖΟΓΛΟΥ
ΠΤΥΧΙΟ ΦΥΣΙΚΗΣ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΜΕΤΕΩΡΟΛΟΓΙΑΣ ΚΑΙ ΦΥΣΙΚΗΣ ΩΚΕΑΝΟΓΡΑΦΙΑΣ

ΑΝΑΠΤΥΞΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΤΕΧΝΙΚΩΝ ΑΦΟΜΟΙΩΣΗΣ ΥΔΡΟΜΕΤΕΩΡΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΤΗΛΕΠΙΣΚΟΠΙΣΗΣ ΣΕ ΜΕΤΕΩΡΟΛΟΓΙΚΑ ΚΑΙ ΕΔΑΦΙΚΑ ΜΟΝΤΕΛΑ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΘΕΣΣΑΛΟΝΙΚΗ 2018

EFTHYMIOS SERPETZOGLOU

DEVELOPMENT AND APPLICATION OF ASSIMILATION TECHNIQUES OF HYDROMETEOROLOGICAL REMOTELY-SENSED DATA IN METEOROLOGICAL AND LAND-SURFACE MODELS

Prepared under the Postgraduate Programme of Meteorology, Climatology and Atmospheric Environment of the School of Geology of Aristotle Univ. of Thessaloniki --- Submitted in the A.U.T. School of Geology in June 2018

Oral Examination Date: 28/08/2018

Annex Number of Scientific Annals of the School of Geology N°: 188

Advisory Committee

Professor Theodoros Karacostas, Supervisor
Associate Professor Prodromos Zanis, Member
Researcher A', Anastasios Papadopoulos, Member

Examining Committee

Professor Theodoros Karacostas
Associate Professor Prodromos Zanis
Researcher A', Anastasios Papadopoulos
Professor Emmanouil Anagnostou
Professor Aristides Bartzokas
Associate Professor Petros Katsafados
Assistant Professor Ioannis Pytharoulis

ΕΥΘΥΜΙΟΣ ΣΕΡΠΕΤΖΟΓΛΟΥ

ΑΝΑΠΤΥΞΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΤΕΧΝΙΚΩΝ ΑΦΟΜΟΙΩΣΗΣ ΥΔΡΟΜΕΤΕΩΡΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΤΗΛΕΠΙΣΚΟΠΙΣΗΣ ΣΕ ΜΕΤΕΩΡΟΛΟΓΙΚΑ ΚΑΙ ΕΔΑΦΙΚΑ ΜΟΝΤΕΛΑ

Εκπονήθηκε στον Τομέα Μετεωρολογίας, Κλιματολογίας και Ατμοσφαιρικού Περιβάλλοντος του Τμήματος Γεωλογίας Α.Π.Θ. --- Υποβλήθηκε στο Τμήμα Γεωλογίας Α.Π.Θ. τον Ιούνιο του 2018

Ημερομηνία Προφορικής Εξέτασης: 28/08/2018

Αριθμός Παραρτήματος Επιστημονικής Επετηρίδας Τμήματος Γεωλογίας N°: 188

Τριμελής Συμβουλευτική Επιτροπή

Καθηγητής Θεόδωρος Καρακώστας, Επιβλέπων
Αναπληρωτής Καθηγητής Πρόδρομος Ζάνης, Μέλος
Ερευνητής Α' Αναστάσιος Παπαδόπουλος, Μέλος

Εξεταστική Επιτροπή

Καθηγητής Θεόδωρος Καρακώστας
Αναπληρωτής Καθηγητής Πρόδρομος Ζάνης
Ερευνητής Α' Αναστάσιος Παπαδόπουλος
Καθηγητής Εμμανουήλ Αναγνώστου
Καθηγητής Αριστείδης Μπαρτζώκας
Αναπληρωτής Καθηγητής Πέτρος Κατσαφάδος
Επίκουρος Καθηγητής Ιωάννης Πυθαρούλης

DEVELOPMENT AND APPLICATION OF ASSIMILATION TECHNIQUES OF HYDROMETEOROLOGICAL REMOTELY-SENSED DATA IN METEOROLOGICAL AND LAND-SURFACE MODELS – *Ph.D. Thesis*

ΑΝΑΠΤΥΞΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΤΕΧΝΙΚΩΝ ΑΦΟΜΟΙΩΣΗΣ ΥΔΡΟΜΕΤΕΩΡΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΤΗΛΕΠΙΣΚΟΠΙΣΗΣ ΣΕ ΜΕΤΕΩΡΟΛΟΓΙΚΑ ΚΑΙ ΕΔΑΦΙΚΑ ΜΟΝΤΕΛΑ – *Διδακτορική Διατριβή*

Citation:
Serpetzoglou E., 2018. – Development and Application of Assimilation techniques of Hydrometeorological remotely-sensed data in meteorological and land-surface models. Ph.D. Thesis, School of Geology, Aristotle University of Thessaloniki, Annex Number of Scientific Annals of the School of Geology No 188 pp. 127

Αναφορά:
Σερπετζόγλου Ε., 2018. – Ανάπτυξη και εφαρμογή τεχνικών αφομοίωσης υδρομετεωρολογικών δεδομένων τηλεπισκόπισης σε μετεωρολογικά και εδαφικά μοντέλα. Διδακτορική Διατριβή, Τμήμα Γεωλογίας Α.Π.Θ., Αριθμός Παραρτήματος Επιστημονικής Επετηρίδας Τμ. Γεωλογίας Νο 188 σελ. 127

ACKNOWLEDGMENTS

This dissertation would not have been possible without the assistance and support of many people who were involved in its preparation and completion. First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Theodore Karacostas, for providing me the opportunity to carry out this research. I am truly thankful for his support, advice and motivation throughout my PhD candidacy. I am further indebted to Professor Emmanouil Anagnostou and Researcher A' Anastasios Papadopoulos for co-supervising my research work. These three people have contributed substantially in my growth as a research scientist and I am really grateful for the knowledge and experience I gained working with them. Their creativity and bright ideas inspired me to conduct novel research and I am truly lucky to have met all three of them.

I am also very thankful to all the Professors at the department of Meteorology and Climatology, School of Geology of AUTH and especially to the Associate Professor Prodromos Zanis and Assistant Professor Ioannis Pytharoulis, who also serve on my dissertation committee, for the valuable discussions during my research progress. At the same time, I would like to express my sincere gratitude to the Professor Aristides Bartzokas and Associate Professor Petros Katsafados for providing valuable feedback while serving at my examining committee, as well as to the researchers and colleagues at the Hellenic Centre of Marine Research, Institute of Inland Waters, which hosted me for significant parts of my research; especially to Dr. Efthymios Nikolopoulos, Dr. Matteo Zampieri and Dr. Viviana Maggioni, who were quite influential in various technical and scientific aspects of this study; with Efthymios we shared not only the same name, but also many fruitful conversations as well as many relaxing breaks at work.

A special reference and expression of gratitude is also well deserved to my former M.Sc. professors at the University of Miami, Dr. Bruce Albrecht and Dr. Pavlos Kollias, whose influence and scientific motivation follow me intact throughout the years.

Finally, I would like to express my deepest gratitude to my family and friends without whom I would not be the person I am now. My heartfelt gratitude goes to my parents, Iordanis and Sofia, for their unconditional love and for constantly pushing me to finish my PhD throughout the years. My wife Katerina further deserves my warmest

acknowledgement for her endless patience and for the strength and encouragement that she was constantly providing me with. My two young daughters, Eleni and Sofia, have also been a source of great inspiration throughout this process as well as a constant reminder that I need to complete the PhD so that they learn the importance of setting and achieving your goals. My brother Christos and my dearest friends and Best men, Georgios Tsiantas, Ilias Patiris and Christos Mandilaras also deserve my special thanks for believing in me and giving me the extra push whenever I needed it.

The author would like to acknowledge the support of the EU Marie Curie Excellence Grant project PreWEC (MEXT-CT-2006-038331).

PREFACE

One of the main objectives of research in the field of hydrometeorology is to advance our understanding of the mechanisms that regulate land-atmosphere interactions and thus improve their mathematical description in the complex schemes of atmospheric and land surface models.

However, parameterization and implicit solutions of the spatiotemporal differential equations that comprise a numerical forecasting model inevitably cause the creation and propagation of inherent errors, which deteriorate the model prediction capability. In addition, an important issue in numerical forecasting is the definition of initial and boundary conditions required for the smooth initiation of a regional model. To address these decades-long limitations, many studies have been carried out towards the improvement of parameterizations of various physical mechanisms and processes. In recent years, however, the abundance of remote sensing data has opened new horizons in the scientific community, directing research on the exploitation of these data through the development and implementation of assimilation techniques in land surface and meteorological models.

In this context, the main objective of the present thesis is the coupling of models and data through the improvement of existing- and the development of new assimilation techniques, which enable more precise determination of initial and boundary conditions and the dynamic correction of numerical model prognostic results. At the same time, the study contributes to the development of new parameterization schemes of soil processes and their interactions with the overlying atmosphere.

These research objectives are achieved on the basis of studying specific cases of special meteorological interest, such as convective thunderstorms and severe floods, which occurred over parts of Europe and North America in the recent past. The study is facilitated by the collection, post-processing and application of hydrometeorological data from ground stations, weather radar and satellites, as well as the application of advanced land surface and meteorological models. The combination of these state-of-the-art tools through data assimilation techniques facilitates an in-depth examination of the interactions between land surface state (soil moisture and temperature, turbulent fluxes etc.) and atmospheric

boundary layer (convection, precipitation, etc.) as well as an improved representation of simulated meteorological fields (precipitation etc.).

The achievements of this study comprise a major theoretical and practical step towards the advancement of integrated systems used already to collect, process, assimilate and forecast hydrometeorological parameters. These systems facilitate significant applications in various areas of social and economic life dependent on accurate and timely weather forecasts, such as protection against natural disasters, management of water resources, agriculture, tourism, and road-, sea- and air transportation.

The originality of the current thesis is established through the publication of much of the embedded work in major peer-reviewed journals (*hyperlinks are provided upon clicking on the underlined text*):

- **Serpetzoglou E.,** E. N. Anagnostou, A. Papadopoulos, E. I. Nikolopoulos, and V. Maggioni, 2010: Error Propagation of Remote Sensing Rainfall Estimates in Soil Moisture Prediction from a Land Surface Model. J. Hydrometeor, 11, 705–720, doi: 10.1175/2009JHM1166.1.
- Zampieri, M., **E. Serpetzoglou**, E. N. Anagnostou, E. I. Nikolopoulos, and A. Papadopoulos, 2012: Improving the representation of river–groundwater interactions in land surface modeling at the regional scale: Observational evidence and parameterization applied in the Community Land Model. J. Hydrol., 420–421, 72-86.
- Papadopoulos, A., **E. Serpetzoglou**, and E. N. Anagnostou, 2008: Improving NWP through radar rainfall-driven land surface parameters: A case study on convective precipitation forecasting. Adv. Water Res., 31, Special Issue on Hydrologic Remote Sensing, 1456-1469.
- **Serpetzoglou, E.**, T. S. Karacostas, I. Pytharoulis, P. Zanis, A. Papadopoulos, and E. N. Anagnostou, 2018: Sensitivity analysis on the ingestion of remotely sensed precipitation data into the land surface scheme of a mesoscale model. *Nat. Hazards Earth Syst. Sci.*, to be submitted.

In addition to the peer-reviewed journals, the author participated in many international conferences and workshops, where the current research work was presented (*hyperlinks are provided upon clicking on the underlined text*):

- **Serpetzoglou E.**, A. Papadopoulos, and E. N. Anagnostou, 2007: Improving Convective Precipitation Forecasting through Assimilation of Data-Forced Land Surface Parameters. Eos Trans. AGU, 88(52), Fall Meeting Suppl., Abstract H42D-04.
- Papadopoulos A., **E. Serpetzoglou**, and E. N. Anagnostou, 2007: The Influence of Assimilating Land Surface Parameters on the Simulation Performance of Warm Season Convective Systems. Proceedings of 9th Plinius Conference on Mediterranean Storms, Varenna, Italy (Plinius Conference Abstracts, Vol. 9, 00036, 2007).
- **Serpetzoglou E.**, A. Papadopoulos, and E. N. Anagnostou, 2008: The response of a NWP model to the ingestion of observed precipitation data on its different land surface schemes in cases of flash-flood inducing storms. EGU General Assembly, Vienna, Austria. (Geophysical Research Abstracts, Vol. 10, 09797, 2008)
- **Serpetzoglou E.**, A. Papadopoulos, and E. N. Anagnostou, 2008: Assimilation of rainfall data into the land surface scheme of a meteorological model and its impact on convective precipitation forecasting. Extended Abstracts, 9th Conference on Meteorology, Climatology and Atmospheric Physics (COMECAP 2008), Thessaloniki, Greece.
- **Serpetzoglou E.**, A. Papadopoulos, and E. N. Anagnostou, 2008: Sensitivity tests on the response of a mesoscale model to the assimilation of observed precipitation data into its land surface scheme. Proceedings of 10th Plinius Conference on Mediterranean Storms, Nicosia, Cyprus (Plinius Conference Abstracts, Vol. 10, 00085, 2008).
- Papadopoulos A., **E. Serpetzoglou**, and E. N. Anagnostou, 2008: Using Satellite Data-Forced Land Surface Parameters to Improve Convective Precipitation Forecasts. 3rd NASA/JAXA International TRMM science conference, Las Vegas, Nevada.
- Papadopoulos A., E. N. Anagnostou, **E. Serpetzoglou**, B. Antonescu, C. Oprea, and S. Burcea, 2008: Investigating the Impact of Lightning Data Assimilation on Flash Flood Forecasting. EGU General Assembly, Vienna, Austria. (Geophysical Research Abstracts, Vol. 10, 07693, 2008)

- E. N. Anagnostou, **E. Serpetzoglou**, , E. I. Nikolopoulos, A. Papadopoulos, Y. Hong, and F. Hossain, 2009: Satellite precipitation error propagation in the prediction of large-scale floods and soil moisture. EGU General Assembly, Vienna, Austria. (Geophysical Research Abstracts, Vol. 11, 09099, 2009)
- Zampieri, M., **E. Serpetzoglou**, E. I. Nikolopoulos, E. N. Anagnostou, and A. Papadopoulos, 2010: Assessing the effect of lateral groundwater fluxes on the simulation of soil moisture in a regional land-surface model. EGU General Assembly, Vienna, Austria. (Geophysical Research Abstracts, Vol. 12, 1213, 2010)
- Zampieri, M., **E. Serpetzoglou**, E. N. Anagnostou, E. I. Nikolopoulos, and A. Papadopoulos, 2011: River-groundwater interactions at the regional scale: evidence from surface stations and satellite observations and a new parameterization in the Community Land Model. EGU General Assembly, Vienna, Austria. (Geophysical Research Abstracts, Vol. 13, 0187, 2011)
- **Serpetzoglou E.**, A. Papadopoulos, E. N. Anagnostou, and T. S. Karacostas, 2009: Studying the Mediterranean land-air interaction feedbacks through long-term assimilation of precipitation data into the WRF land surface schemes. 3rd HyMeX workshop, 1-4 June 2009, Heraklion (Gournes), Crete, Greece.

TABLE OF CONTENTS

Chapter 1 Introduction

1.1 Hydrometeorology

Hydrometeorology is the science that focuses on all aspects of the water cycle in the earth's planetary system. It lies in the margins of hydrology and meteorology and focuses primarily on the water state exchange processes that occur between the upper planetary layer and lower atmosphere. These processes can be divided in two main categories: soil water processes (groundwater flow, infiltration, surface runoff, discharge into oceans) and atmospheric boundary layer processes (evaporation, transpiration, water transport in the air, condensation, precipitation). These processes interact with each other through the interchange of the water state (liquid – gas – solid), which is also facilitated by planetary factors (e.g. solar heating, topography etc.), thus forming the complex earth's hydrometeorological system depicted in Fig. 1.

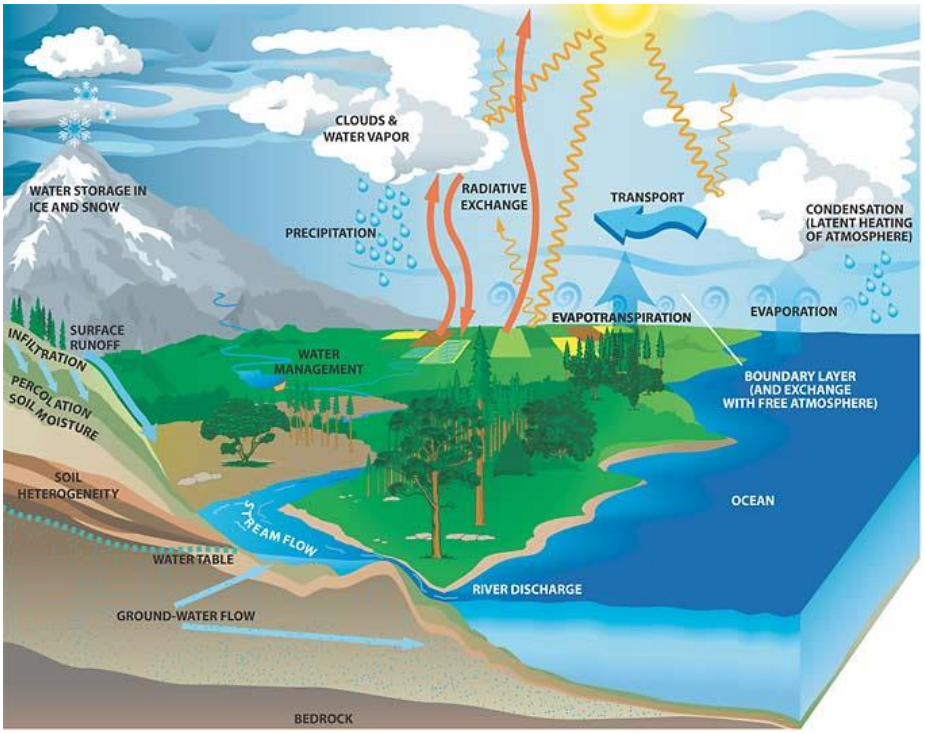


Figure 1.1. The planetary water cycle (picture is from https://smap.jpl.nasa.gov/resources/48/the-water-cycle/)

1.2 Land-surface interactions at regional and global scale

1.2.1 Data sources and modelling systems

Land surface state predictability has been the focus of increasing scientific interest over the past few decades, mainly due to the significant role and complex nature of land-atmosphere interactions that greatly impact weather and climate. Key land surface variables, such as soil moisture and soil temperature, greatly influence the atmospheric boundary layer at both short-range and seasonal time scales (e.g., Beljaars et al. 1996; Betts 2004; Fischer et al. 2007). Accurate prediction of land surface states can be advantageous to weather and regional/seasonal forecasts that notably depend on initial and boundary land surface conditions (e.g., Koster and Suarez 2001; Drusch and Viterbo 2007; Papadopoulos et al. 2008). The level of this accuracy is highly variable and pertains to the variety of sources and methods used to define the land surface conditions. Land surface models (LSMs), coupled, or not, to atmospheric models, and forced by in-situ or remotely-sensed hydrometeorological data, constitute the main tool today for the prediction of land surface parameters, suitable for use in the initialization and boundary conditions updating of advanced weather and regional climate models (e.g., Robock et al. 2003; Koster et al. 2004a; Rodell et al. 2005; Maggioni and Houser 2017). The verification and uncertainty characterization in these land data assimilation systems (LDAS) becomes of increasing importance, especially under the consideration of the latest advances on remote sensing of near-surface soil moisture and other important land surface variables (Entekhabi et al. 2004; Walker and Houser 2004; Reichle et al. 2004; 2007; 2008; 2013; Entekhabi et al. 2010).

Until recently, direct observations of land surface variables have been generally based on point measurements of soil moisture and energy fluxes from in-situ stations, which are limited to a few locations around the globe either under the auspices of specific field experiments or – less frequently – as part of long-term in-situ networks (Robock et al. 2000). Measurement networks of in situ sensors (such as USDA's Soil Climate Analysis Network (SCAN) or NOAA's Climate Reference Network (CRN) in the continental United States) have potentially high soil moisture measurement accuracy but are spatially

very sparse. Moreover, primary efforts to adequately capture soil moisture properties from space-borne microwave sensors (e.g., Walker and Houser 2004; Reichle et al. 2004) faced significant technical challenges and long delays (Leese et al. 2001).

However, the last decade has been thriving with advancements in soil moisture retrievals as a direct result of the constantly growing scientific interest on soil moisture properties and effects combined with significant technological progress (De Jeu and Dorigo 2016). Research algorithms that were initially applied to sensors designed for other purposes, e.g., for measuring wind speed [e.g., the Advanced Scatterometer (ASCAT)], sea ice, or atmospheric parameters [e.g., the TRMM Microwave Imager (TMI) and the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E)], have developed into fully operational soil moisture products (Dorigo and De Jeu 2016; see this publication also for a thorough reference-based review of the major applications of satellite-observed soil moisture). These research activities were the prelude for the first multi-satellite soil moisture dataset released under the auspices of the ESA's Water Cycle Multi-mission Observation Strategy (WACMOS) and the subsequent Climate Change Initiative (CCI) projects (Liu et al. 2012; 2011b; see also http://cci.esa.int/projects). More specifically, the soil moisture CCI project (http://www.esa-soilmoisture-cci.org/) produced a complete and consistent multi-decadal record (currently from 1978 to 2014) through the combination of various soil moisture products from active and passive microwave sensors, which allows to look at long-term variability and change in observed global-scale soil moisture (already recognized as an Essential Climate Variable (ECV) since 2010) and to correlate long-term soil moisture dynamics with the variability observed in other ECVs. Fang et al. (2016) intercompare the various CCI products with LSM and in-situ soil moisture measurements over the continental US for a quite long period (2000-2013), showing that the merged CCI product is capturing soil moisture properties very well.

Further to the aforementioned projects and datasets, the next-generation comprehensive global monitoring of soil moisture has been achieved and is ongoing with the use of advanced satellite sensors as part of two major missions: ESA’s Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al. 2001; Mecklenburg et al. 2012) and NASA’s Soil Moisture Active Passive (SMAP) mission (Entekhabi et al. 2010). By globally monitoring surface and near-surface soil moisture over land at a coarser (i.e., ~40

km for SMOS data) or finer (i.e., ~10 km for SMAP data) spatial scale and at a temporal resolution of 2-3 days, these two missions are expected to cover a wide area of applications and thus advance science in meteorology, climatology, agriculture and hydrology/water resources management. SMOS data have recently been incorporated into the soil moisture CCI dataset, and together with the high-resolution SMAP time-series, already contribute to weather and seasonal climate forecasting through the use of soil moisture data as model initial and boundary conditions and/or assimilated data (for more information, please visit https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/smos/ as well as consult the SMAP Handbook, Entekhabi et al. 2014, available online at https://smap.jpl.nasa.gov/mission/description/). Recent studies have been already proving the significance of assimilating the prescribed datasets into LSMs. Soil moisture simulations from the Noah LSM, either coupled to the Weather Research and Forecasting (WRF) model (Lin et al. 2017) or as part of the NASA LIS (Blankenship et al. 2016), were characterized by notably improved error metrics when assimilating SMOS soil moisture data.

However, despite their extreme importance, the SMOS and SMAP time series face technical setbacks (e.g., coarse spatial, temporal and sensing depth resolution of SMOS data; coarse temporal resolution and limited duration of the SMAP mission). As a result, these satellite sources cannot fully provide global soil moisture estimates with high spatial and temporal resolution, optimal sensing depth, and desired accuracy over moderate vegetation conditions. Moreover, McNally et al. 2016 claim that the SMOS and SMAP data are still too short to conduct studies that demonstrate the utility of these data for operational applications, or to provide historical context for extreme wet or dry events. Taking all that into account, LDAS still remain the major tool to provide global estimates of soil moisture and other land surface variables that can be directly assimilated into weather- and climate-scale models for retrospective studies or forecasting applications (e.g., Walker et al. 2003). The typical structure of a LDAS is depicted in Fig. 1.1.

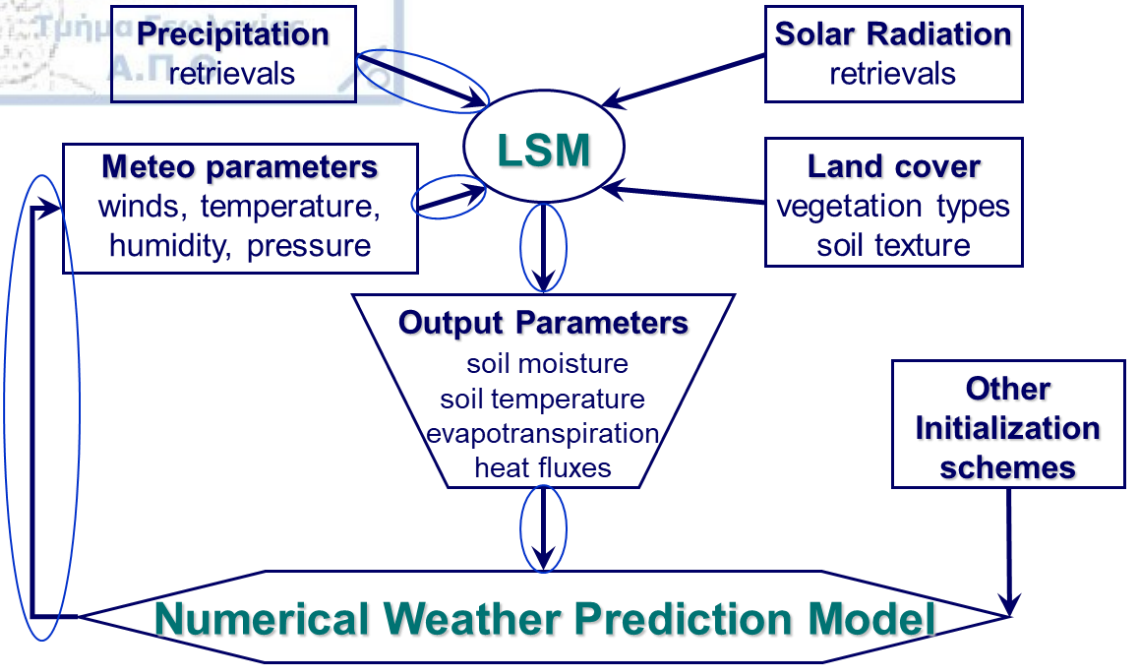


Figure 1.2. Typical Structure of a Land Data Assimilation System

The Land Information System (LIS; Kumar et al. 2006) is currently the major software framework for high performance terrestrial hydrology modelling and data assimilation developed with the goal of integrating satellite and ground-based observational data products and advanced modelling techniques to produce optimal fields of land surface states and fluxes (source: http://lis.gsfc.nasa.gov/). The North American LDAS (NLDAS; Mitchell et al. 2004; see also Schaake et al. 2004), the Global LDAS (GLDAS; Rodell et al. 2004) and the Famine Early Warning Systems Network (FEWS NET) LDAS (FLDAS) are the three main instances of LIS, each one utilizing multiple LSMs and multiple sources of satellite and ground-based hydrometeorological observations to provide optimal fields of land surface states and fluxes in near-real time (for more information, please visit http://ldas.gsfc.nasa.gov/).

Amongst all hydrometeorological variables used to force the various LDAS, precipitation receives the greatest attention. Accurate precipitation measurements are critical for the implementation of the LDAS near-real time simulations (Gottschalck et al. 2005, Liu et al. 2011a). The requirement for global or extensive regional simulation coverage can be facilitated either with the use of global/regional climate model rainfall

outputs or the exploitation of advanced satellite rainfall retrievals. Both pathways have advantages and disadvantages, yet the on-going development of high-resolution (<0.5 degree and <3-hourly) global rainfall estimates from a combination of infrared (IR) and passive microwave (PMW) retrievals (e.g., Sorooshian et al. 2000; Joyce et al. 2004; Huffman et al. 2007) has driven attention to remotely-sensed precipitation estimation. Such datasets have been under thorough investigation and intercomparison to define the optimal rainfall estimates based on regional and time-scale criteria (e.g., Ebert et al 2007; Tian et al. 2007; Hossain and Huffman 2008; Anagnostou et al. 2010; Gehne et al. 2016). Based on its successful predecessor (e.g., the Tropical Rainfall Measurement Mission; TRMM), the newly-launched Global Precipitation Measurement (GPM) mission (Smith et al. 2007, Skofronick-Jackson et al. 2017) is designed to facilitate this effort of providing high-resolution global rainfall estimates from a deployed constellation of satellite-based passive microwave sensors. By providing more accurate estimates of the rate of transfer of water from the atmosphere to the surface, GPM reduces a significant source of uncertainty in the global water/energy budget. GPM-era precipitation observations are expected to enhance the accuracy and spatio-temporal resolution of the precipitation forcing in LDAS, leading to improvements in the prediction of land surface states. Scientists will combine GPM observations with land surface data to provide better estimates for soil moisture, temperature, and snowpack, which leads to better predictions of vegetation cover, improved weather forecasts and integrated hydrologic models (Kirschbaum et al. 2017, Skofronick-Jackson et al. 2017).

With the GPM mission being assessed in terms of societal values, there is a need of identifying and prioritizing hydrometeorological uses of the current data sources and methods to make this mission as effective as possible over land. One of the key challenges of the GPM mission and the projects thoroughly described here is the improvement of numerical weather prediction, especially the accurate forecasting of rapidly developing mesoscale convective systems (MCS), which are usually responsible for the heaviest and most destructive rainfall and flood events in major continental regimes. To moderate these hazards, systems have been developed for issuing short-range (up to 48 h) quantitative precipitation forecasts (QPF) based on numerical weather prediction (NWP) models and observational inputs (e.g., synoptic observations, surface and remotely sensed data). An

important component of QPF is the simulation of deep moist convection processes and the consequential surface rainfall. At the resolutions typically used by regional NWP (grid increments ranging from 5 to 20 km), numerical models exhibit low skill at forecasting highly variable convective precipitation events. The main error sources can be considered the (i) lack of sufficient data to specify initial and boundary conditions, (ii) coarse grid resolutions that limit our ability to correctly specify those features acting to trigger convection in model initialization (e.g., Kain and Fritsch 1992) and (iii) weak assumptions used in developing the convective parameterization schemes (e.g., Mellor and Yamada 1982). Arguably, the high non-linearity in modelling atmospheric processes causes sensitivities to even small perturbation in the atmospheric initial state, the quality of which dictates the accuracy of QPF.

1.2.2 Interaction between soil moisture and precipitation

One avenue for improving the convective precipitation forecasting skill based on remotely sensed data is through a better estimation of the energy and moisture exchanges between land surface and the atmosphere (Papadopoulos et al. 2008). As discussed earlier, soil moisture is known to play an important role in the partitioning of energy between sensible and latent heat fluxes as well as in driving the total moisture budget and temperature variations of the lower atmosphere (e.g., Clark and Arritt 1995; Betts 2004; Fischer et al. 2007). Furthermore, it has long been debated whether there is a direct effect of evapotranspiration originating from soil moisture abundance on precipitation amount (Schär et al. 1999). When the details on the water recycling rate were clarified and it was recognized that water particles travel long distances before returning again to the earth's surface, studies (e.g., Budyko 1974) criticized the earlier assumptions that precipitation primarily derives from local evapotranspiration. However, the impact of initial soil state conditions in weather forecasting models was identified by several modelling studies in the 80s and 90s (e.g., Mahfouf 1991; Viterbo 1995), while further studies have shown strong sensitivity of precipitation predictions on evapotranspiration disturbances. Beljaars et al. (1996) found that perturbations in the initial soil-moisture settings of the European

Centre for Medium-Range Weather Forecasts (ECMWF) land-surface scheme resulted in significant differences in the modelled precipitation 2 to 3 days later. Ramos da Silva and Avissar (2006), using the Regional Atmospheric Modelling System (RAMS) to simulate the evolution of convection in a deforested region of the Amazon basin, confirmed that higher soil moisture not only produces more rainfall but also delays its formation. Drusch and Viterbo (2007) showed that observations of soil moisture are capable of improving the turbulent surface fluxes and consequently the weather forecast on large geographical domains.

Similar interactions at seasonal to sub-seasonal scales have also been verified; Koster and Suarez (2001) found that soil moisture is an important source of forecast skill for the predictability of precipitation, and Zhang and Frederiksen (2003) showed that the initial soil moisture conditions supplied to an atmospheric general circulation model (AGCM) affect both temperature and precipitation forecasts. Although Schär et al (1999) and Koster et al (2004b) argued that the increase of precipitation with increasing soil moisture depends heavily on model, regional and seasonal differences, still they show that there are areas around the world that exhibit strong coupling between soil moisture and precipitation. By testing initializations of varying soil moisture in a regional climate model over Europe, Schär et al (1999) found evidence of an indirect mechanism of precipitation dependence on soil moisture, according to which the surplus of precipitation is transported over long distances by atmospheric circulation, but the efficiency of precipitation processes is controlled by the soil state conditions. Further, in a pioneering study that utilized a synthesis of ensemble forecasts originating from an extensive suite of AGCMs, Koster et al. (2004b) showed that it is the transition zones between wet and dry climates where soil moisture has higher probability of influencing precipitation. In the same study such an area appears to be the central Great Plains of North America, which is the focus of our study.

The above studies have demonstrated that soil moisture strongly influences the boundary layer dynamics. However, the role of soil temperature in the evolution of the lower atmosphere should not be underestimated. For example, outgoing longwave radiation is a function of soil temperature that directly affects the surface radiation budget. Further, the ground heat flux depends on soil temperature and soil moisture conditions, as

well as vegetation coverage, atmospheric conditions, and the thermophysical properties of the soil. Apparently, the characterization of the spatial and temporal variability of soil moisture and soil temperature fields is a critical factor in formulating and modifying the mesoscale atmospheric circulations.

1.2.3 Hydrological processes at regional scale

The other primary aspect affecting the spatial and temporal variability of soil moisture and soil temperature fields involves processes that take place below the earth's surface (Zampieri et al. 2012). Groundwater is a basic component of the hydrosphere and plays a fundamental role in many processes that affect the atmosphere and the biosphere. Groundwater sustains streams, lakes, wetlands and the related ecosystems (Alley et al. 2002, Dahm et al. 2003). It provides a lower boundary condition for soil moisture and a direct source of water for plant roots, thus affecting evapotranspiration, especially in warm periods and shallow water table conditions (Schmidhalter et al. 1994, Snyder and Williams 2000, Scott et al. 2006, Steinwand et al. 2006, Yeh and Famiglietti 2009, Xie and Yuan 2010), and the carbon cycle (Ju et al. 2006). Soil moisture, in turn, affects surface temperature through the control on the partitioning of sensible and latent heat flux (Zampieri et al. 2009).

Groundwater is implicitly accounted in the land surface schemes of many climate models for the computation of surface and sub-surface runoff (Koster et al. 2000, Ducharne et al. 2000, Walko et al. 2000, Chen and Kumar 2001, Seuffert et al 2002, Gedney and Cox 2003, Yang and Niu 2003, Niu and Yang 2003, Niu et al. 2005). A number of these models follow the TOPMODEL approach (Beven and Kirkby 1979) assuming that, for each grid cell of the model, ponded areas exist because of the interaction of groundwater dynamics and subgrid orography. Therefore, runoff production is expected to increase in case of shallow water table. The TOPMODEL can be considered as a one-way surface-groundwater interaction model, which accounts for the flux of water from the saturated zone to the surface. Many studies have explicitly accounted for the groundwater to improve the soil moisture and evapotranspiration representation in the land surface schemes for the general circulation models and regional climate models (Famiglietti and Wood 1994,

Stieglitz et al. 1997, Gutowski et al. 2002, York et al. 2002, Liang et al. 2003, Maxwell and Miller, 2005, Yeh and Eltahir 2005a,b, Cohen et al. 2006, Niu et al. 2007, Miguez-Macho et al. 2007, Anyah et al. 2008, Yuan et al 2008, Jiang et al. 2009, see also Fan et al. 2007 for a review). These studies have shown that incorporating the water table dynamics enhances modelled evapotranspiration and eventually reduces bias in the simulated precipitation, especially in the warm season of humid and semi-humid climates, as for instance monsoon-influenced climates. In fact, in regions with shallow water table, groundwater can determine the soil moisture profile and provide a direct source of water for transpiration, thus increasing the "memory" of soil conditions to precipitation and enhancing the persistence of intraseasonal and interannual precipitation in regional climate models, as suggested by Bierkens and van der Hurk (2007).

An important source of groundwater is reinfiltrated water from streams (Sophocleous 2002). This process can be simulated explicitly at the river scale (Osman and Bruen 2002), and at the watershed scale through high-resolution coupled land surface – groundwater models (Kollet and Maxwell 2006). It is usually neglected in regional land surface models (LSMs) because of the lack of resolution that is needed to simulate the local infiltration of river water and the lateral hydrological processes along river corridors. As noted by Zhang and Montgomery (1994), a good representation of these lateral processes requires a spatial scale of at least 10 m resolution while a representative LSM scale in regional climate applications is 10-30 km. River transport models (RTMs) are used in some LSMs to simulate fresh water fluxes into the oceans that is needed to close the global water budget. However, the interaction of river water with groundwater is generally neglected in these models.

Studies have shown that rivers and stream-groundwater interactions are responsible for the lagged correlation between precipitation over the mountains and wetter soils over the planes (Kingston et al. 2009, Wedgbrow et al. 2002). Consequently, lack of proper representation of these processes in the climate models could result in biases in the simulated surface climate. Furthermore, stream-groundwater dynamics can alter the spatial variability in the soil moisture field that could influence the local atmospheric circulation and moist convection (Weaver 2006, Steiner et al. 2009) with feedback effects on the surface climate itself. Unfortunately, a direct validation of LSMs in terms of comparison

with observed groundwater or soil moisture data is difficult, because data are sparse and representative of very small areas relative to a typical LSM grid resolution. However, changes in the hydrological cycle could be investigated through the indirect effect on surface temperature, for which satellite products exist at the desired resolution and spatial coverage. This is an aspect we explore further in this dissertation.

1.3 Dissertation outline

The first part of the dissertation (chapter 2; Serpetzoglou et al. 2010) deals with the part of the continental water cycle that involves the land surface state predictability. It includes the study of the propagation of error of remotely-sensed precipitation data to the soil moisture fields simulated by an advanced LSM. The study facilitates the in-depth investigation and comparison of two major error sources in the simulation of land surface state properties, namely the error induced by rainfall forcing and the error induced by model internal parameterizations. Through this error sensitivity analysis, two other major contributions are achieved; namely, the comparison of the performance of three extensive sources of remotely-sensed rainfall data (i.e., NEXRAD, TRMM and CMORPH) as well as the evaluation of the performance of the advanced LSM (i.e., the Community Land Model, version 3.5), all of which will further be utilized in the following chapters for improving model parameterizations and developing new data assimilation techniques.

The second part of the dissertation (chapter 3; Zampieri et al. 2012) investigates the part of the continental water cycle that involves underground processes, namely the interaction between water reservoirs over land (e.g., rivers, lakes etc.) and the underlying groundwater. The study uses the same advanced LSM exploited in chapter 2 and contributes significantly to the hydrological science through the improvement of the model parameterization of the interactions between surface- and groundwater. With the present study, CLM3.5 accounts not only for the groundwater effect on lakes and rivers but also for the opposite effect of surface water reservoir discharge into groundwater at the regional scale.

Through the seasonal examination of the simulated soil moisture fields, both studies discussed in chapters 2 and 3 manage to identify the climatological effects of the prescribed

analyses. Specifically, the CLM error statistics and the new river-groundwater interaction parameterization highly depend on the wetness conditions that prevail during the simulation period.

The third part of the dissertation (chapter 4; Papadopoulos et al. 2008, Serpetzoglou et al. 2018 – to be submitted) covers the part of the water cycle that involves the interactions between the upper soil layer and the overlying atmospheric boundary layer. The chapter discusses the development and application of a new assimilation technique through the ingestion of remotely-sensed precipitation data in the land surface schemes of advanced mesoscale models. The technique significantly improves the models' quantitative precipitation forecasting capability in cases of extreme thunderstorms, thus offering a valuable tool to meteorologists worldwide. The assimilation technique is tested for two different advanced weather forecasting systems (i.e., the POSEIDON and the WRF systems) in two different continental regimes (i.e., continental USA and Europe, respectively) with two different sources of remotely-sensed rainfall for the LSM data ingestion (i.e., NEXRAD and CMORPH estimates, respectively). Moreover, an extensive feedback investigation is performed, advancing our understanding of the complex land-atmosphere interaction processes.

The linkage between all three studies discussed in the present dissertation is the worldwide need to mitigate the effects of extreme weather phenomena, and especially of highly-dangerous thunderstorms produced by mesoscale convective systems, the precipitation quantities of which combined with specific topographic domains and river networks often cause significant human losses and huge infrastructure damages.

Chapter 2 Error propagation study in land surface modelling

2.1 Introduction

Remote sensing of rainfall from satellite data sources will always be subject to error that can be of complex structure at high spatio-temporal scale (Hossain and Anagnostou 2004; 2006). The propagation of this error through the non-linear land-atmosphere interaction processes resolved by LSM can impact soil moisture prediction in a way that depends on scale, precipitation error characteristics and the complexity of modelling system. This stresses the need for thorough and systematic investigations and quantification of the error propagation properties for optimal LDAS simulations (Maggioni et al. 2012). A key feature for the accurate definition of soil moisture prediction uncertainty is the quantification of interactions between the rainfall forcing uncertainty with the LSM parametric error (Hossain and Anagnostou 2005a; 2005b, Maggioni et al. 2011, Gottschalk et al. 2005). Gottschalk et al. (2005) compared multiple satellite, model and ground-based rainfall datasets, which were then used to force year-long GLDAS simulations to perform qualitative diagnosis of their impacts on land surface states. They found that GLDAS land surface states are sensitive to different precipitation forcing; percent differences in volumetric soil water content (SWC) between simulations ranged from 75% to 100% for both summer and winter seasons, and these differences were generally 25%–75% less than the percent precipitation differences, indicating that GLDAS and specifically the Mosaic LSM acted to generally “damp” precipitation differences. However, evident were also areas where the changes in SWC were equivalent to the precipitation changes. Hossain and Anagnostou (2005a) were the first to explore the issue of the complex interaction between rainfall and modelling uncertainties in LSM soil moisture prediction. They performed numerical experiments using ensemble-based techniques to isolate and characterize propagation of errors in the satellite rainfall estimation alone, the LSM parametric uncertainty alone (manifesting as non-uniqueness in soil hydraulic parameters), and the combined data-modelling uncertainty. They found that the contribution of precipitation error was generally lower than that of modelling uncertainty, with satellite retrieval error

contributing between 20% and 60% of the total uncertainty in soil moisture prediction in the cases of modelling accuracy ranging from low to high, respectively.

In this chapter, we expand the work by Hossain and Anagnostou (2005a) seeking to examine and quantify the two sources of uncertainty in the simulation of soil moisture fields from an offline LSM (member of the GLDAS/LIS) forced by three different sources of remotely-sensed precipitation estimates, namely two satellite and one from ground radar (Serpetzoglou et al. 2010). The assessment of both sources of error in soil moisture prediction is uniquely facilitated using in-situ measurements of soil moisture, rainfall and other meteorological variables on a small domain in the Midwestern US capturing the State of Oklahoma. The region is covered by a dense network of environmental monitoring stations named Oklahoma Mesoscale Network (Mesonet; Brock et al. 1995) available over a long-term period (1997-present). The abundance of measurements over the Mesonet region in combination with the climatic characteristics of this area (e.g. standing in the transition zone between wet and dry areas in the United States; Koster et al. 2004b) made this region suitable for pursuing the goals of our study. The Mesonet station data are used here as reference for evaluating the remotely sensed rainfall retrievals at high spatio-temporal scales as well as the LSM simulations of soil moisture. The study also presents a rigorous benchmarking of the Mesonet network as to its accuracy in deriving area rainfall estimates at the resolution of satellite products (0.25 degree/3-hourly) based on comparisons against the most definitive Micronet station measurements (see section 2.2 for more details on the Mesonet and Micronet networks).

2.2 Study area and datasets

The current error propagation study was facilitated with the use of data from various sources. First, surface meteorological observations (e.g., pressure, air temperature, relative humidity, solar radiation, wind speed, and rainfall) at 5-min resolution as well as soil moisture observations at 3 depths (5, 25 and 60 cm) at 30-min resolution were provided from 115 stations of the Mesonet network (Fig. 2.1). Within the Mesonet network, the U.S.D.A. Agricultural Research Service's (ARS) Grazinglands Research Laboratory (GRL) has established a smaller-area yet denser network of hydrometeorological stations

(42 stations from 1994 to 2005, 20 core stations thereafter) that monitors the environmental conditions of the Little Washita watershed, called the Little Washita Micronet. Three stations in the Oklahoma Mesonet are located in the northeast, south, and west areas of the watershed (Fig. 2.1). Both datasets are quality controlled and flagged for bad quality data. Data were available for the period 1997-2006 for the Mesonet stations, and the period 2002-2004 for the Micronet network.

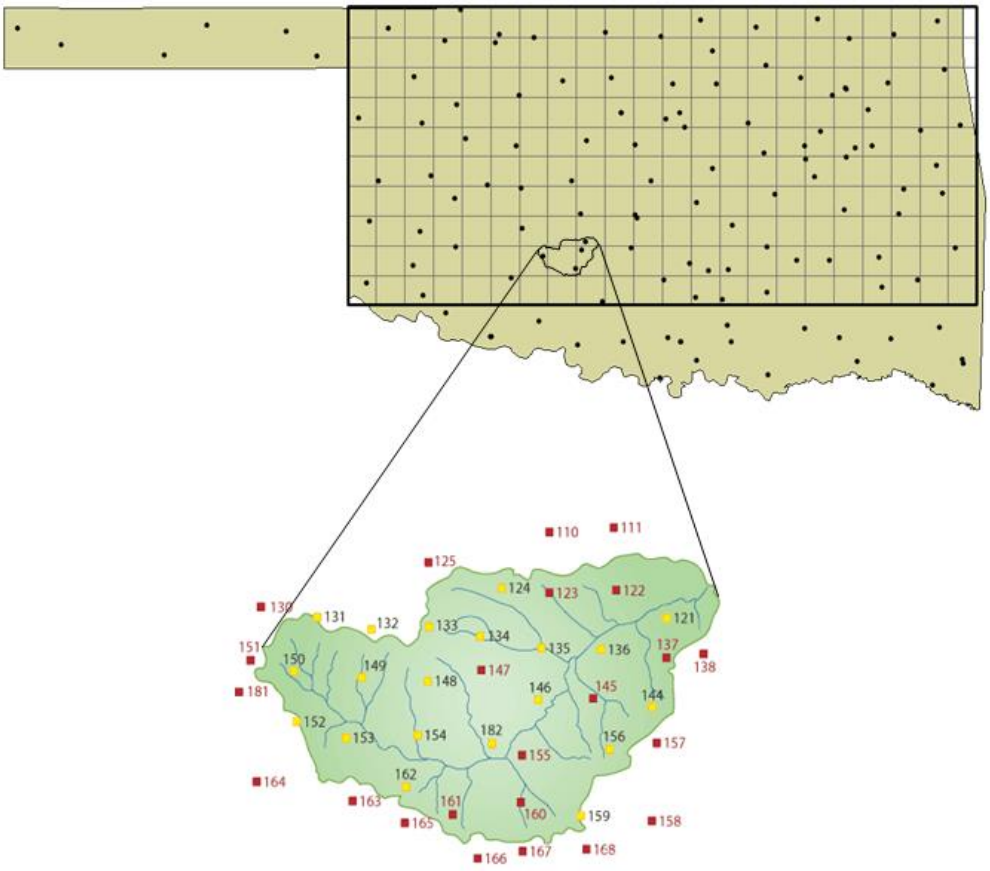


Figure 2.1. Schematic of the Oklahoma Mesonet and Micronet networks. The locations of the Mesonet stations are indicated by the black dots. The centres of the cells of the 0.25-degree grid shown in the same graph represent the locations where soil moisture data are extracted based on the application of gage interpolation technique. The ARS Little Washita River watershed is also expanded in a bigger diagram showing the locations of the Micronet stations (picture obtained from http://ars.mesonet.org/sites/). The stations marked with red colour are retired since 2005.

Further, three different remotely-sensed rainfall datasets were used for the 3-year period 2004-2006: the National Aeronautics and Space Administration (NASA) Goddard

Space Flight Centre's (GSFC) Tropical Rainfall Measuring Mission (TRMM) 3B42 product (obtained at 0.25-degree/3-hr spatio-temporal resolution), the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Centre's (CPC) morphing technique (CMORPH) product (obtained at 8-km/30-min spatio-temporal resolution) and the Stage IV US WSR-88D radar network (NEXRAD) estimates (obtained at 4-km/1-hr spatio-temporal resolution). Hereafter, the three datasets will be referred to as TRMM, CMORPH and NEXRAD, respectively.

Measurements from satellite passive microwave sensors are the primary source for the precipitation estimates of both CMORPH and TRMM rainfall products. Multiple satellite platforms are used to facilitate maximum coverage and enhanced temporal sampling for both products, including TRMM, Defense Meteorological Satellite Program (DMSP), NOAA, and Earth Observing system (EOS) platforms. However, the way that these sensors are inter-calibrated to extract each product differs; In 3B42, passive microwave (MW) sensors are inter-calibrated to TRMM's combined precipitation radar (PR) and Microwave Imager (TMI) retrievals, while CMORPH uses TMI and DMSP Special Sensor Microwave Imager (SSM/I) as calibration reference, with TMI having the highest precedence whenever available (Tian et al. 2007). Both datasets use IR data from geostationary satellites to fill in MW coverage gaps, yet in different ways. For the CMORPH product, the dynamic morphological characteristics (such as shape and intensity) of the precipitation features are morphed at consecutive times between MW sensor samples by performing a time-weighted linear interpolation. This process yields spatially and temporally continuous passive MW rainfall fields that have been guided by IR imagery and yet is independent of any infrared temperature-based inversion to rainfall rate (Joyce et al. 2004). The TRMM 3B42 product, on the other hand, uses MW-calibrated IR precipitation estimates directly, to fill the MW coverage gaps. A further difference between the two products pertains to the use of surface gauge measurement information; CMORPH only uses satellite estimates, while TRMM 3B42 combines the merged MW- and IR-based estimates with gauge observations via scaling of the individual 3B42 3-hourly precipitation values to monthly gauge analysis. A detailed comparison of these two products in terms of their potential use in LDAS simulations was performed in Tian et al. (2007), who infer that the TRMM 3B42 product is suitable for long-term, retrospective, and climatological studies due to its

reduced biases on longer time scales, while CMORPH is recommended for short-term applications due to its higher probability of detection of rainfall events. The radar rainfall fields used in this study were extracted from the Stage IV National Weather Service (NWS) precipitation estimation algorithm product that involves real-time adjustment of the radar rainfall estimates based on mean-field radar-rain gauge hourly accumulation comparisons and merging of hourly radar with gauge-interpolated rainfall fields (Fulton et al. 1998; Lin and Mitchell 2005).

The NEXRAD and CMORPH rainfall products were rescaled to the TRMM spatio-temporal resolution (0.25-deg×0.25-deg, 3-hourly) such that common spatial grid (depicted in Fig. 2.1) and temporal scale are used for all products to allow for direct comparisons and model input implementation. The in-situ Mesonet meteorological data (including the rainfall measurements) were also interpolated to this common spatial grid using the inverse distance weighting (IDW) technique and averaged to 3-hr increments. The Kriging interpolation scheme was preferred due to the longer spatial correlation of soil moisture to create the same grid for the Mesonet soil moisture data. A point to note is that both interpolation techniques influence the characteristics of rainfall forcing and soil moisture measurements, mainly in terms of not entirely preserving the rainfall peaks and soil moisture maxima, and thus creating spatially smoother fields. The reader is referred to Anagnostou et al. (2010) for a detailed analysis of the error properties of the aforementioned rainfall products used in this study.

2.3 Land surface model and experimental setup

The NCAR Community Land Model (CLM, version 3.5) is used in this study to simulate the land surface and land-air exchange processes. CLM3.5 is a well-documented model (see http://www.cgd.ucar.edu/tss/clm/distribution/clm3.5/index.html) that has been designed to integrate all land processes into a single modelling system. It is one of the models used in GLDAS (Rodell et al. 2004) and LIS (Kumar et al. 2006) systems and receives extensive attention in many land surface modelling studies (e.g., Oleson et al. 2008, Stöckli et al. 2008, Tian et al. 2008).

The model components comprise biogeophysics (i.e., surface fluxes of energy, moisture, and momentum), hydrologic cycle, biogeochemistry and dynamic vegetation (for the purposes of the current study we utilized the biogeophysics and hydrology components only). Based on externally-provided atmospheric forcing data (e.g., precipitation, radiation, wind speed, air temperature, and humidity fields) the model computes a number of prognostic surface variables that include runoff, soil moisture and temperature in various soil layers, water intercepted on the canopy, leaf temperature, latent and sensible heat fluxes. CLM3.5 has one vegetation layer, like most land surface models, ten unevenly spaced vertical soil layers (the respective depths are defined at 0.7, 2.8, 6.2, 11.9, 21.2, 36.6, 62.0, 103.8, 172.8, and 286.4 cm) with variable hydraulic conductivity, and up to five snow layers depending on the total snow depth (Oleson et al. 2004). The land surface is represented by 5 primary sub-grid land cover types (glacier, lake, wetland, urban, vegetated) for each grid cell. The vegetated portion of a grid cell is further divided into patches of plant functional types (PFTs), each with its own leaf and stem area index and canopy height. Each sub-grid land cover type and PFT patch is a separate column for energy and water calculations at every time step. The high-resolution surface datasets currently used in CLM3.5 are based on newly-developed Moderate Resolution Imaging Spectroradiometer (MODIS) products (Lawrence and Chase 2007; see Table 2.1 for the Oklahoma region). Most surface processes such as evaporation from the ground, transpiration from the plants' rooting zone, soil and snow water propagation, leaf temperature and fluxes, soil and snow temperature, and phase change are parameterized through physical equations. The parameterization of runoff-related processes is based on the TOPMODEL concept (Beven and Kirkby 1979; Niu et al. 2005).

Table 2.1. High resolution surface datasets used in version 3.5 of CLM.

Surface Field	Resolution	Aggregation Method
Percent glacier	0.05º	Area average
Percent lake	0.05º	Area average
Percent wetland	0.05º	Area average
Percent sand and clay	0.05º	Soil mapping unit with greatest areal extent in grid cell
Soil colour	0.05º	Soil colour class with greatest areal extent in grid cell
PFTs (percent of vegetated land)	0.05º	Area average
Monthly leaf and stem area index	0.05º	Area average
Canopy height (top, bottom)	0.05º	Area average

Combining CLM3.5 and the datasets described earlier, various experiments were designed. The control experiment included the use of the post-processed Mesonet meteorological observations as input in CLM3.5 (named CLM-MESONET), while three other experiments were facilitated with the use of the three remotely-sensed rain products as rainfall input in place of the respective Mesonet observations (named CLM-TRMM, CLM-CMORPH and CLM-NEXRAD, respectively). A half-hourly time step was chosen for all CLM3.5 runs to provide for high-accuracy simulations (with a constant rainfall input for each half hour during a 3-hr period), while the model outputs were 3-hourly to match the satellite rainfall products time sampling. All simulations were 3-year long, starting on January 1st 2004, 00 UTC and ending on December 31st 2006, and were initialized with the CLM3.5 output on January 1st 2004, 00 UTC produced by an earlier-in-time 7-year run (1997-2003). This spin-up time is effectively longer in duration than common practice for land surface models (Cosgrove et al. 2003b).

Each experiment outputs various soil and energy properties, and our study is focused on the soil moisture fields. It is worth mentioning that a quadratic fitting model was applied to those fields to account for the discrepancy in depth with the actual Mesonet observations. Specifically, soil moisture output at the first four model soil depths (i.e., 0.7, 2.8, 6.2, 11.9 cm) was taken into account to provide the best available estimate for soil moisture at 5 cm, based on a 2nd order fitting equation defined uniquely for each grid cell

at every time step. The specific design of these four experiments allows for an in-depth analysis of the propagation of the rainfall-product error through the soil moisture prediction as well as the relationship between data and modelling uncertainties. The results presented in this study focus on the summer periods (i.e., June-July-August; JJA) of 2004 and 2006.

2.4 Results

2.4.1 Assessment of Mesonet sampling on rainfall estimates and soil moisture simulations

The high-resolution Micronet rainfall dataset facilitates an in-depth examination of the coarser-resolution Mesonet network as to its adequacy to provide accurate precipitation forcing data for the land surface model. A Mesonet benchmarking experiment was conducted with CLM3.5 using grid-cell average rainfall from the dense Micronet network as a reference to force a cell in the Mesonet domain shown in Fig. 2.1 (this is named CLM-MICRONET experiment). Fig. 2.2 shows a comparison between Mesonet and Micronet in terms of the grid-cell average rain rates (at 3-hourly time intervals) and the corresponding model-predicted near-surface (5-cm depth) soil moisture values for the 2004 summer season. As noted from the figure, the Micronet average values of rainfall accumulation are very close to the rainfall accumulation values from the respective Mesonet grid cell, which leads to almost identical temporal evolutions of CLM-MESONET and CLM-MICRONET soil moisture values (see Fig. 2.2a). The discrepancies observed in Fig. 2.2a between model results and soil moisture observations by Mesonet stations (i.e., the fact that model results seem to be much more responsive than the observations in terms of response to rainfall events and dry-down) is attributed to two main sources: (i) errors due to modelling caused by incorrect model parameters and errors in the *assumed-close-to-the-truth* reference rainfall and (ii) errors in the soil moisture observations caused by sensor measurement uncertainties and the smoothing effect of the Kriging interpolation scheme applied to the measurements to create the 0.25-degree grid cell averages. Unfortunately, the errors in soil moisture observations, which are not quantifiable from currently available data, are lumped

into the "modelling" error and do not affect the "rainfall-induced" error. Consequently, significant soil moisture observing errors would skew the relationship between "modelling" and "rainfall-induced" land surface model errors, therefore, results presented in this study should be viewed as conservative in terms of the modelling uncertainty.

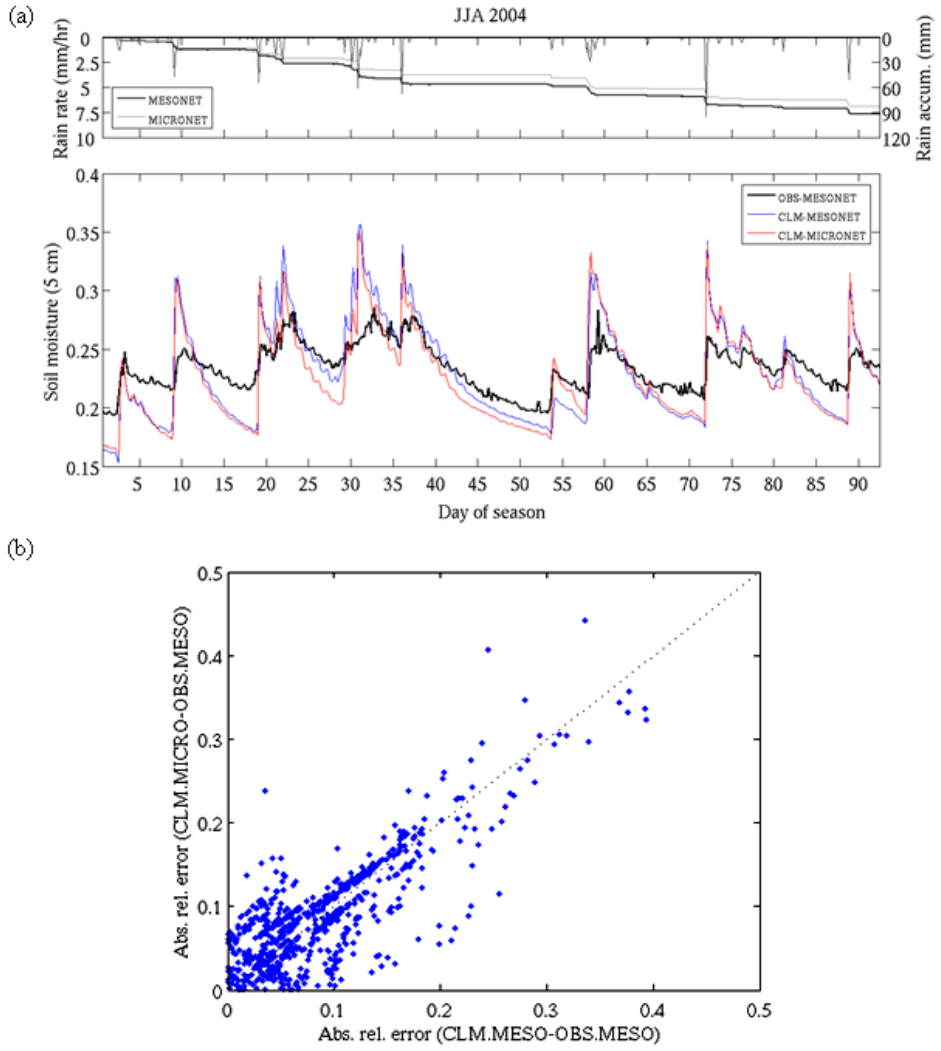


Figure 2.2. Comparison between high-resolution (Micronet) and low-resolution (Mesonet) rainfall forcing for the 2004 warm season. (a) Upper panel: Rain accumulation (mm) from the Micronet and Mesonet networks and rain rate (mm/hr) from the respective Mesonet/Micronet grid cell. Lower panel: Time series of observed near-surface soil moisture (Mesonet) and predicted near-surface soil moisture values (CLM-MESONET and CLM-MICRONET experiments) for the Mesonet/Micronet grid cell. (b) Absolute relative errors in near-surface soil moisture for CLM-MICRONET versus the respective error for CLM-MICRONET, for each 3-hr record of the Mesonet/Micronet grid cell.

The discussion that follows offers a quantitative statistical evaluation of the error properties of rainfall products and soil moisture estimates qualitatively captured in Fig. 2.2a. As expected, the correlation coefficient between rainfall estimates of Micronet and Mesonet is very high (0.95). However, the Mesonet probability of rain detection (POD) relative to Micronet is about 63%. Although this may seem a low detection score, the majority of the non-detected Micronet rainfall values is below 0.05 mm/hr, as revealed by the conditional POD (at the 0.05 mm/hr threshold) that is equal to 90%. On the other hand, the false alarm rate (FA) is very low in both cases (7% for the unconditional case and 6% for the conditional case). The reader is also urged to look at Anagnostou et al. (2010) for an in-depth statistical analysis of the same Mesonet and Micronet rainfall datasets. Regarding the CLM3.5 soil moisture estimates, the correlation coefficient between CLM-MESONET and CLM-MICRONET is even higher than the respective rainfall value (0.97). Both model estimates are also highly correlated to the Mesonet soil moisture observations (0.88 for CLM-MESONET and 0.84 for CLM-MICRONET), although here the respective bias and error values are quite high due to the spatial interpolation of the initial point measurements. The absolute relative errors of these model estimates (defined as absolute relative differences from the respective 2004-summer-season 3-hourly Mesonet soil moisture observations) are directly compared in Fig. 2.2b exhibiting very similar characteristics.

Moreover, the relative Root Mean Square Error (RMSE) of rainfall (defined as the RMS of the difference between Mesonet and Micronet rainfall estimates normalized by the mean value of Micronet rainfall, conditional to Micronet rainfall > 0.05 mm/hr) is 0.49 as opposed to a value of 0.07 for the relative RMSE of soil moisture (defined as the RMS of the differences between CLM-MESONET and CLM-MICRONET soil moisture estimates normalized by the mean value of CLM-MICRONET soil moisture). Further, the conditional mean biases are 1.58 for rainfall (Mesonet to Micronet rainfall estimates) and 1.02 for soil moisture (CLM-MESONET to CLM-MICRONET soil moisture estimates). These values indicate a “dampening” of the relatively significant rainfall error through the land surface model simulation, verifying similar qualitative (Gottschalck et al. 2005) and quantitative (Hossain and Anagnostou 2005a) results presented in earlier studies. Another aspect that highlights the importance of the latter inference pertains to the dampening of

error in rainfall associated with the application of the IDW interpolation technique. The foregoing analysis thus supports our notion that the low-resolution Mesonet observations (as compared to the high-resolution Micronet dataset) are sufficient for studies that include forcing of land surface models and respective error propagation quantification.

2.4.2 Rainfall-induced vs. modelling-induced errors in soil moisture prediction

Fig. 2.3 depicts the 2004 and 2006 warm season temporal evolution of observed area-average rain rates and accumulation from the whole Mesonet grid and both observed (Mesonet) and predicted (CLM3.5 experiments) mean near-surface (5-cm depth) soil moisture values. In both warm seasons we observe a noteworthy correspondence between rain occurrence and changes in both observed and predicted near-surface soil moisture. Enhanced rainfall is always followed by significant increase in near-surface soil moisture, which is yet expected but still quite prominent.

Comparing the Mesonet-observed to the model-predicted area-average near-surface soil moisture time series, we notice some differences between the two summer periods under study. In summer 2004, there is good agreement between measured and predicted soil moisture in the study region, especially with regard to CLM-MESONET, CLM-TRMM and CLM-NEXRAD experiments. However, there are periods when these three experiments seem to overestimate soil moisture as compared to the Mesonet measurements, which coincide with short-term peaks in soil moisture magnitude following events of intense rainfall. As discussed before, part of this overestimation of the CLM3.5 simulation outputs may be due to the underrepresentation of soil moisture maxima from the Mesonet interpolated fields and errors due to modelling. The CLM-CMORPH experiment does not perform as well as the experiments based on the other two remote sensing rainfall products, since the respective time series is characterized by significant bias with respect to the measured soil moisture. This is mainly due to an actual overestimation in summer season precipitation by the CMORPH product (described in Anagnostou et al. 2010), which is consistent with results of previous studies (e.g., Tian et al. 2007).

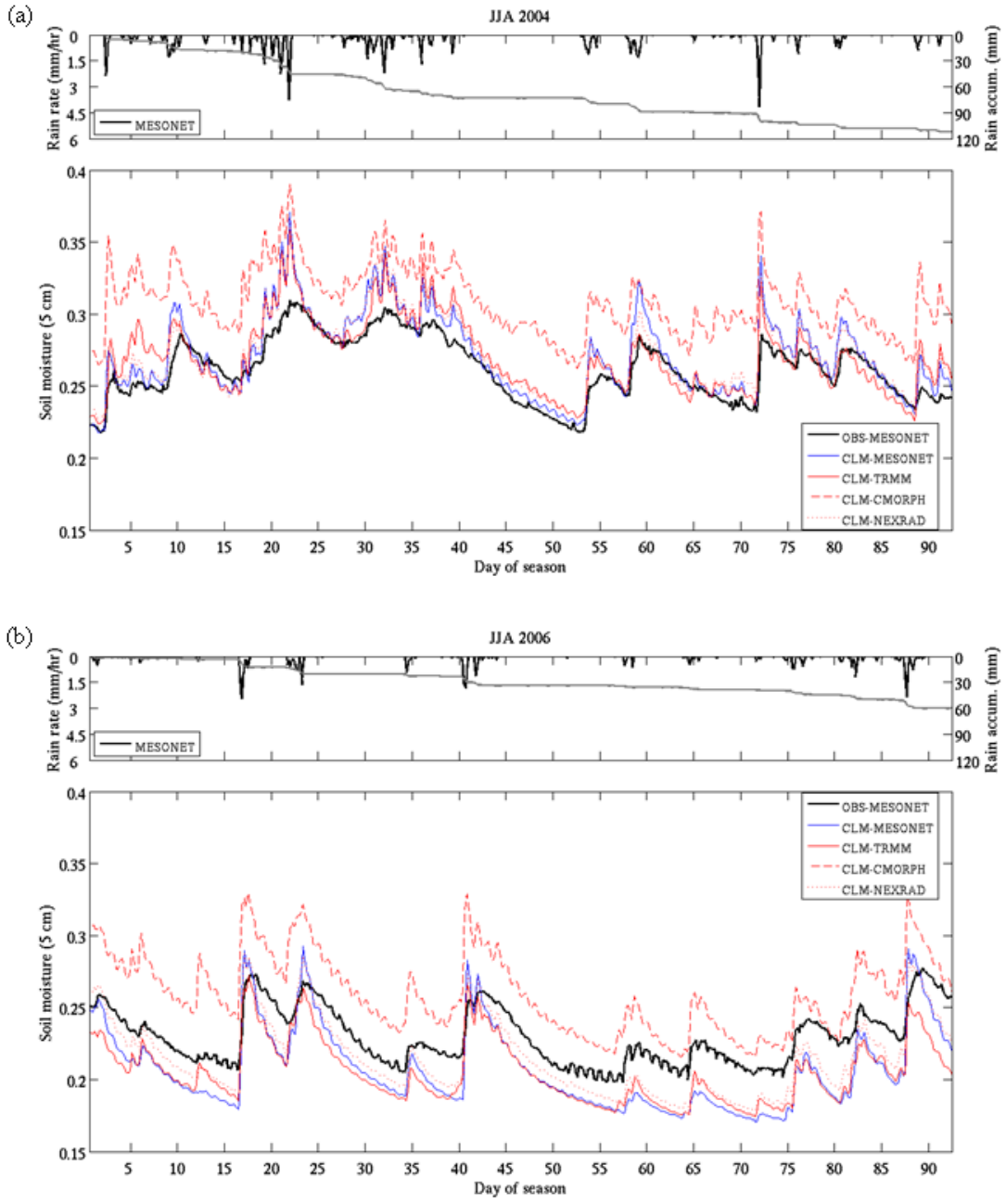


Figure 2.3. Time series of observed rain rate (mm/hr), rain accumulation (mm) and soil moisture (at 5 cm) from the Mesonet network and predicted soil moisture from the four CLM3.5 experiments, for the summer periods of 2004 (a) and 2006 (b). In all cases, the average value of the gridded domain shown in **Figure 2.1** is depicted.

The 2006 summer period is generally drier than 2004, as indicated by the accumulated rainfall amounts and the soil moisture magnitudes shown in Fig. 2.3b. Here, CLM-MESONET, CLM-TRMM and CLM-NEXRAD experiments tend to slightly underestimate soil moisture magnitudes most of the time (periods with no or almost no rainfall). A closer look at the soil moisture temporal evolution during the early days of June 2006 (Fig. 2.3b) reveals that a portion of this bias originates from the preceding period (e.g., the spring season – not depicted here). However, this propagated spring-induced bias is steadily increasing with time as the simulation advances to July and August. On the other hand, these CLM3.5 simulations seem to agree with the Mesonet measurements during the peak periods of heavy rainfall and increased soil moisture, although the latter should be mainly ascribed to the non-preservation of high-frequency modulations in interpolated soil moisture that relate to short-term (convective) heavy rainfall events. CLM-CMORPH again exhibits significant positive bias with respect to the measured soil moisture.

Further, one could argue that the observed underestimation of the area-average near-surface soil moisture magnitude by most CLM3.5 simulations during summer 2006 could be attributed to the long-term and continuous nature of the simulations. If such an argument was true, we would expect to observe a respective – lower in magnitude but still evident – bias in the summer 2005. However, this is not the case here, as the 2005 warm season is characterized by a good correspondence between measured and model-predicted soil moisture (except for the CLM-CMORPH experiment), very similar to what we observed in Fig. 2.3a for the 2004 warm season (graph not shown here). It is worth mentioning that the summer 2005 was also a relatively wet period for Oklahoma region with a Mesonet summer rainfall accumulation of about 105 mm, a value very close to the 2004 summer period rainfall accumulation indicated in Fig. 2.3a.

A point to note from the preceding analysis is that the model's effectiveness to represent the soil moisture temporal evolution seems to depend on climatological factors; except for CMORPH, model experiments seem to perform very well in the case of the wet summer of 2004 (and 2005), while the performance skill is lowered during the dry summer of 2006. To investigate the wetness effect on the error characteristics of remotely-sensed rainfall-forced soil moisture predictions we introduce a rainfall climatology parameter, ΔR, defined as

$$\Delta R_i = \frac{R_i - R_{mean}}{R_{mean}},\ i = 1,...,220 \quad \textit{(Equation 2.1)}$$

where R_i is the Mesonet rainfall for the ith grid cell averaged over the entire 3-month period and R_{mean} is the mean value for the entire 10x22 grid area and the 3-month period. ΔR can be interpreted as a climatological wetness indicator of the area covered from the respective grid cell; positive (negative) values of ΔR would indicate areas that are generally moist (dry) with respect to the climatology of the entire Oklahoma region – defined as the 3-month average rainfall value. Subsequently, (i) the RMS of modelling error defined as the difference of grid-cell near-surface soil moisture between the CLM-MESONET prediction and the respective Mesonet measurement and (ii) the RMS of rainfall-induced soil moisture error for the three remotely sensed rainfall products (defined as the difference in grid-cell near-surface soil moisture between CLM-MESONET prediction and CLM-TRMM, CLM-CMORPH and CLM-NEXRAD predictions, respectively) are calculated and plotted against ΔR. The plots for the warm seasons of 2004 and 2006 are shown in Fig. 2.4 and Fig. 2.5, respectively. Indeed, in summer 2004 we observe a distinct pattern of rainfall-induced soil moisture error discrepancy between positive and negative values of ΔR (see Figs. 2.4b, 2.4c and 2.4d); i.e., the rainfall-induced soil moisture RMSE seems to be lower in the moister areas of Oklahoma in all three rainfall products. This behaviour, though, does not characterize the modelling error pattern (Fig. 2.4a). On the other hand, a uniform pattern with respect to dry and moist areas is observed at all error plots of the warm season 2006 (Fig. 2.5), and this is attributed to the dry conditions that prevailed in Oklahoma during summer 2006, as opposed to the summer of 2004. Thus, it is inferred that the abundance (lack) of rainfall that is physically translated to increased (decreased) values of near-surface soil moisture influences the spatio-temporal characteristics of the land surface model's representation of soil moisture. The above discussion indicates that the characteristics of soil moisture prediction error, particularly due to the contribution of the remotely-sensed rainfall-forcing uncertainty, are dependent on the wetness conditions defined by the spatio-temporal climatology of rainfall.

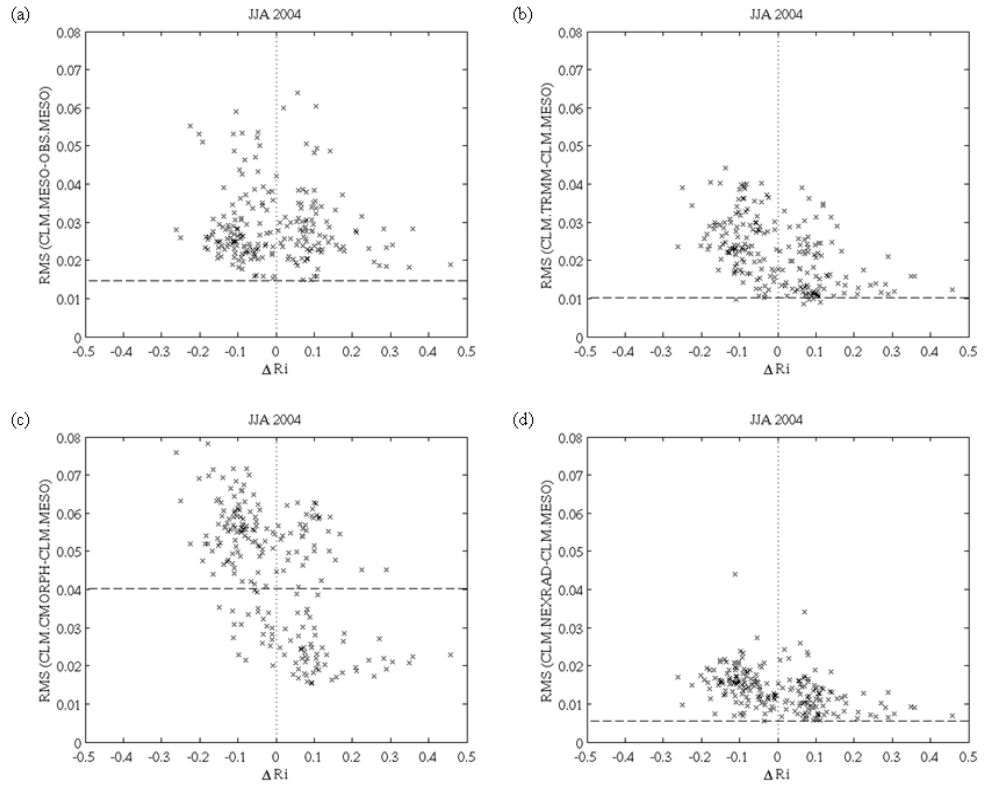


Figure 2.4. Near-surface soil moisture grid-cell RMSEs vs. rain climatology (ΔR_i) for (a) CLM-MESONET, (b) CLM-TRMM, (c) CLM-CMORPH and (d) CLM-NEXRAD, for the summer period 2004. The dashed lines indicate the RMSE of the respective mean values (averages for the entire gridded domain), which is independent of ΔR_i.

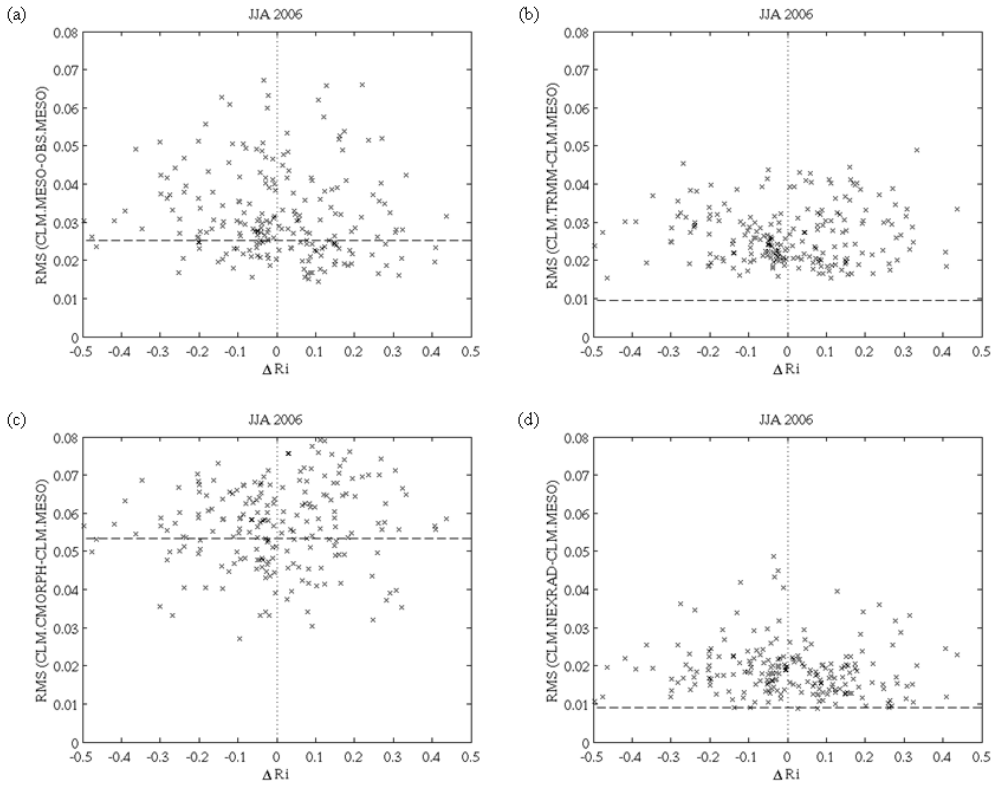


Figure 2.5. As in **Figure 2.4**, but for summer 2006.

A further difference between the two summer seasons pertains to the error magnitude. A cross-examination of Figs. 2.4 and 2.5 with respect to the area-average modelling- and rainfall-induced errors reveals that both types of error are lower in 2004 than 2006 (except for the case of the TRMM-forced RMSE, which is about 0.01 for both summer seasons). The area-average modelling RMSE in 2004 is approximately 0.015, as opposed to an almost double value (0.026) for 2006. The relatively low NEXRAD-forced (0.0055) and relatively high CMORPH-forced (0.04) RMSEs in summer 2004 also appear significantly increased in summer 2006 (0.009 and 0.053, respectively). These differences in RMSEs could be attributed to the increased bias characterizing the model-predicted soil moisture during the warm season of 2006 (discussed earlier), and not be entirely associated with the drier conditions that prevailed over the same period.

Another noteworthy feature observed in Figs. 2.4 and 2.5 is the relatively similar values of modelling- and rainfall-induced errors in soil moisture prediction. If the CLM-

CMORPH experiment is excluded, which is characterized by large values of grid-cell RMSEs both in 2004 and 2006, CLM-TRMM and CLM-NEXRAD rainfall-induced errors are of the same order with the modelling error (below 0.03). This is examined more in depth in Figs. 2.6 and 2.7, where the rainfall-forced soil moisture RMSE is plotted against the model-induced soil moisture RMSE for all grid cells for the warm seasons of 2004 and 2006, respectively. These grid-cell distributions reveal that for most areas in Oklahoma the 2004 CLM-TRMM rainfall-induced error is quite similar to the modelling error, and the respective CLM-NEXRAD error is slightly lower than the modelling error. In the 2006 summer season, the grid-cell CLM-TRMM and CLM-NEXRAD distributions show similar variability, with a tendency towards higher modelling errors. Clear differences emerge after averaging up to the domain scale; Both CLM-TRMM and CLM-NEXRAD average rainfall-induced errors are lower than the respective model-induced error in summer 2006, as opposed to similar or slightly different errors in summer 2004 (differences are shown to be statistically significant at the 99% confidence interval). CLM-CMORPH rainfall-induced error, however, is much higher than the modelling error in soil moisture prediction for most areas of the Oklahoma domain and in both seasons under study, which is yet associated with the prescribed CMORPH rainfall bias manifestation during the warm period of the year. It is noted here that the modelling error estimates are expected to be slightly overestimated, since they are subject to the source of uncertainty associated with the Mesonet soil moisture observation and spatial interpolation errors. However, our results and conclusions are not affected, since this uncertainty has similar effect on the modelling error of both seasons under comparison in the current study.

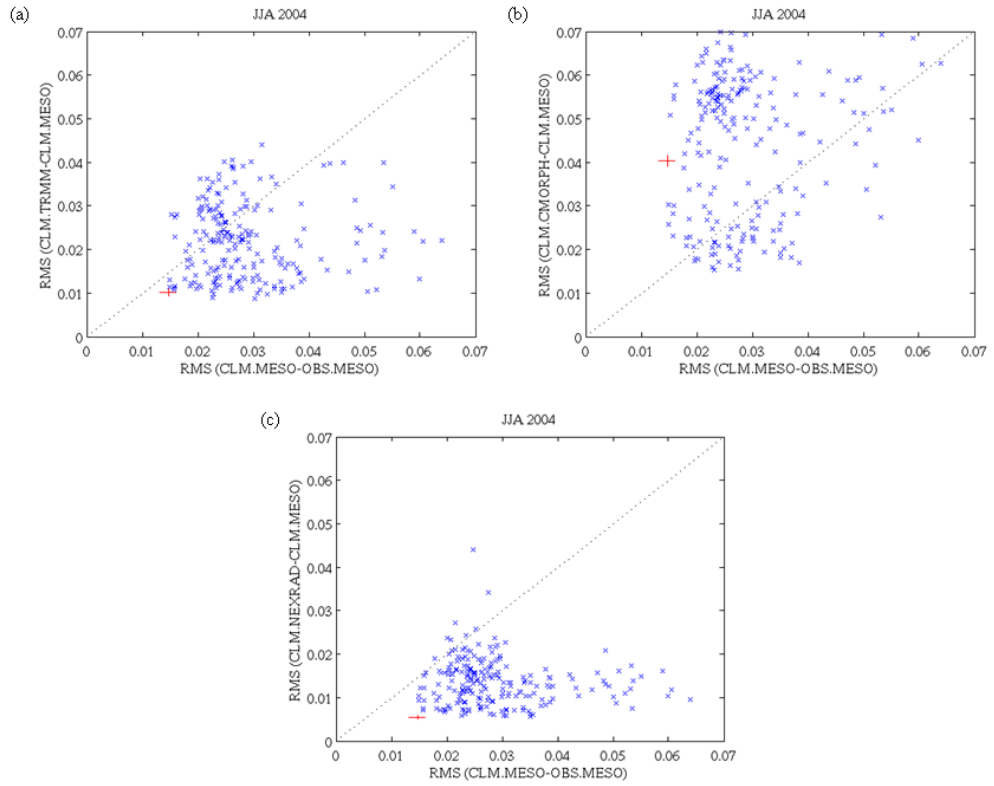


Figure 2.6. Near-surface soil moisture RMSE of (a) TRMM-forced, (b) CMORPH-forced and (c) NEXRAD-forced experiment (e.g., rainfall-induced error) vs. near-surface soil moisture RMSE of Mesonet-forced experiment (e.g., model-induced error) for each grid cell, for the 2004 summer period. The red marker denotes the relationship between the RMSEs of the respective mean values (averages for the entire gridded domain). The length of each red bar represents the respective 99% confidence interval.

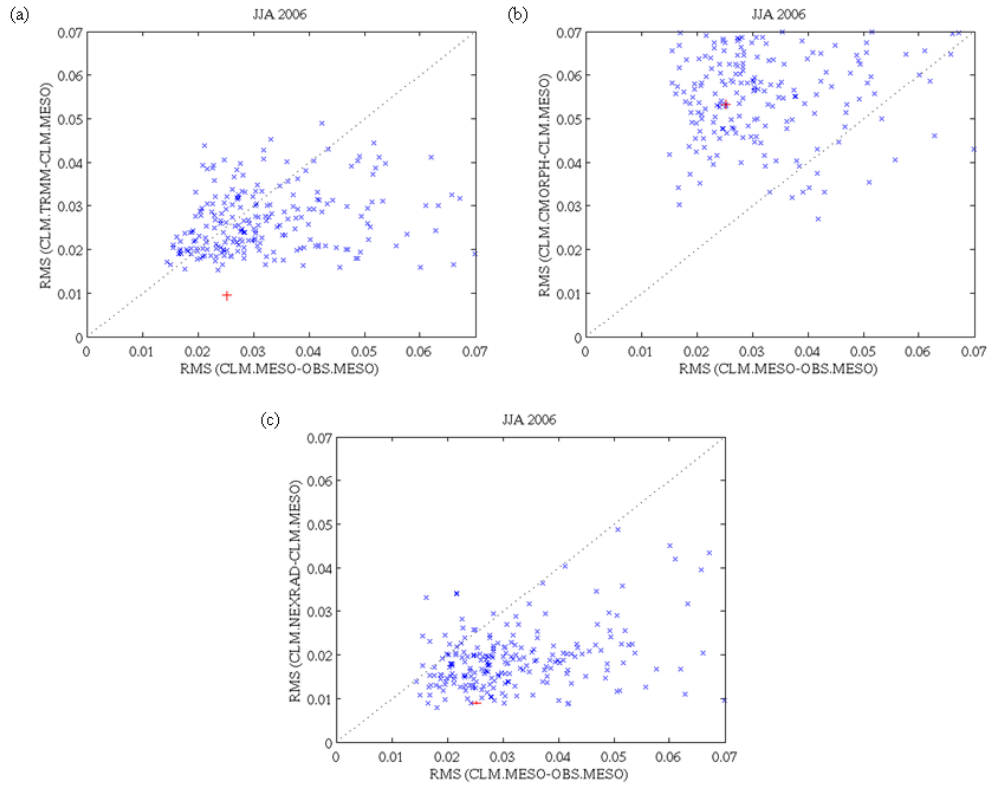


Figure 2.7. As in **Figure 2.6**, but for summer 2006.

The results discussed above are generally consistent with results presented in Fig. 2.8 in Hossain and Anagnostou (2005a) that were based on numerical experiments performed with ensemble-based techniques (see section 2.1). Direct quantitative comparisons between the results of the two studies cannot be performed, since different statistical approaches are used. Hossain and Anagnostou (2005a) addressed the issue of the relative contributions of modelling and rainfall uncertainties to the total uncertainty, which we have not directly tackled in the current study. However, both studies isolate and inter-compare modelling and rainfall uncertainties. Specifically, Fig. 2.8 in Hossain and Anagnostou (2005a) indicates that the modelling and rainfall uncertainties are approximately of the same magnitude when the modelling accuracy is high, as opposed to the case of low modelling accuracy that is characterized by a rainfall error contribution much lower than the modelling error. Our results are equivalent in the sense that the moist 2004 summer season represents our physically-based case of higher modelling accuracy, which is

characterized by almost similar values of domain-averaged modelling and rainfall contributed errors in the soil moisture simulations (TRMM and NEXRAD experiments; Figs. 2.6a and 2.6c, respectively), and the dry 2006 summer season corresponds to the case of lower modelling efficiency, demonstrating lower TRMM and NEXRAD domain-scale rainfall-induced errors in soil moisture simulations than the actual domain-scale modelling error (Figs. 2.7a and 2.7c, respectively).

2.4.3 Effect of scale on rainfall-to-soil moisture error propagation

An important addition to the issues addressed so far is the investigation of the effect of spatial scale on the rainfall error propagation in soil moisture prediction. The horizontal grid of 0.25°x0.25° used in this study is the high-resolution grid utilized by most space-borne rainfall retrievals. However, past remotely-sensed precipitation estimates as well as global model rainfall outputs use coarser spatial resolutions (see for example the datasets used in Gottschalck et al. 2005 and Ebert et al. 2007). Thus, it is important to examine the rainfall-induced error impact on the current LSM simulations of soil moisture rescaled to coarser horizontal grids. For that reason, the initial grid of 0.25°x0.25° was post-aggregated to larger-scale domains, e.g. 0.5°x0.5°, 1°x1°, 1.25°x2.75° and 2.5°x2.75°. All measured rain rates (from Mesonet, TRMM, CMORPH and NEXRAD) as well as near-surface soil moisture (as measured from Mesonet and as predicted from the four CLM3.5 experiments) were interpolated to the aforementioned spatial scales. A point to note is that the methodology applied here imposes some limitations in the sense that no actual simulations were performed at the coarser scales under consideration. The resolution effect on modelling is a major issue that merits detailed investigation that is beyond the scope of this study.

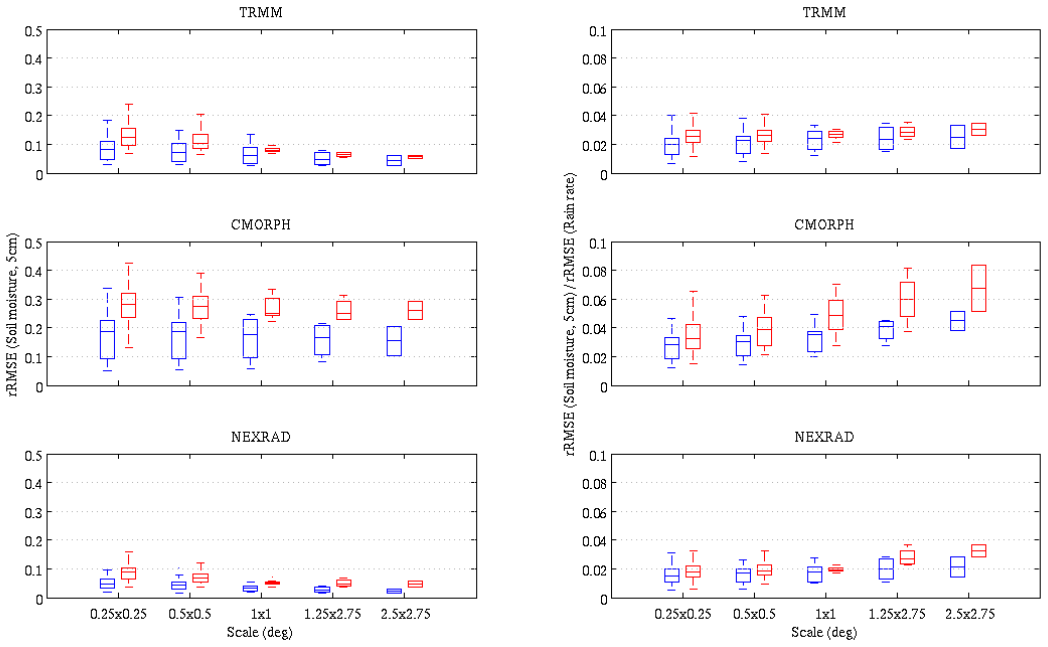


Figure 2.8. Box and whisker plots indicating the scale effect on rainfall error propagation in soil moisture prediction for the warm seasons of 2004 (blue colours) and 2006 (red colours). Left panels show the relative rainfall-induced RMSE in near-surface soil moisture and right panels the ratio of this relative RMSE (rRMSE) over the relative RMSE in rain rate, for CLM-TRMM (upper panels), CLM-CMORPH (middle panels) and CLM-NEXRAD (lower panels) for each of the prescribed scales. The boxes have lines at the lower and upper quartile values and the median value, while the whiskers are the dashed lines extending from each end of the boxes showing the extent of the rest of the data.

Fig. 2.8 depicts the scale dependence of the relative rainfall forcing RMSE in soil moisture prediction as well as the ratio of this relative RMSE to the respective relative rain rate RMSE, for each one of the three remote-sensing products and respective model experiments and for both seasons under study. This ratio is an objective measure of the propagation of the actual error in rainfall estimates (model input error) to the soil moisture estimates (model output error) through the CLM3.5 simulations. The scale effect is evident in all cases. With respect to the rainfall-induced uncertainty in soil moisture alone (Fig. 2.8, left panels), we observe a decreasing trend with coarser spatial resolutions for all three experiments and for both seasons that could be attributed to the smoothing effect intrinsically associated with the aggregate nature of larger-scale fields. However, if this uncertainty is normalized by the respective error in rain rate (Fig. 2.8, right panels), then we clearly notice an increasing propagation error as the spatial resolution is getting lower

for all three remotely-sensed rainfall products and respective CLM3.5 experiments. The latter result further implies that the error in rainfall estimation decreases with coarser spatial resolutions, and actually at a rate higher than the decreasing rate of the rainfall-induced uncertainty in soil moisture mentioned above. We expect that this gradient would be even stronger if land surface modelling would be performed at the respective resolutions due to non-linear hydrological processes. The fact that the rainfall propagation error grows with coarser spatial resolution exemplifies the need for using higher spatial resolution rainfall retrievals in the prediction of hydrological variables and underlines the importance of the GPM mission objectives.

Breaking the results of Fig. 2.8 down to each precipitation product and respective CLM3.5 experiment, we note high consistency with results presented in previous subsections. NEXRAD-related errors and error propagation ratios are the lowest observed for all scales under consideration (below 10% and 4%, respectively, for both summer seasons), and the distribution for each scale is practically uniform around the median value. The respective TRMM-related error properties are similar with the exception of slightly higher values (below 15% and 4%, respectively). CMORPH-related errors and error propagation ratios are much higher (at the order of 20-30% and up to 8%, respectively), while the results are skewed towards the lower quartile values (especially in the case of summer 2004). Between the two different years, the summer season 2004 exhibits the lowest error magnitudes and error propagation ratios in accordance with the aforementioned positive effect of moisture abundance on rainfall-induced soil moisture uncertainty.

2.5 Chapter summary and discussion

The current chapter presented an in-depth investigation of the properties of remotely-sensed rainfall error propagation in the prediction of near-surface soil moisture from a land surface model. A plethora of data, including in-situ measurements and remotely-sensed retrievals, used either as forcing or reference for LSM simulations, facilitated a detailed analysis of the interaction between rainfall-induced uncertainty and modelling uncertainty and their impact on land surface state predictability. The study examined and quantified

both sources of uncertainty in the simulation of soil moisture fields from an offline LSM (CLM3.5) forced by three different sources of remotely-sensed precipitation estimates; two satellite (TRMM, CMORPH) and one from ground radar (NEXRAD). The assessment of both sources of error in soil moisture prediction was performed for the warm seasons of 2004 and 2006 and through the use of in-situ measurements of soil moisture, rainfall and other meteorological variables on a small domain in the Midwestern US capturing the State of Oklahoma (Oklahoma Mesonet). The study also presented a rigorous benchmarking of the Mesonet network as to its accuracy in deriving area rainfall estimates at the resolution of satellite products (0.25 degree/ 3-hourly) based on comparisons against the most definitive measurements of a smaller yet denser network in Southwestern Oklahoma (Micronet).

Our results indicated a dependence of the CLM3.5 efficiency in predicting near-surface (at 5-cm depth) soil moisture fields on the rainfall spatio-temporal climatology; model experiments showed better performance in the case of the relatively moist summer 2004, as opposed to a lowered performance skill during the relatively dry summer 2006. Further, moister areas within the Oklahoma region were associated with reduced rainfall-based error with respect to drier areas. All error magnitudes (e.g. modelling and remotely-sensed rainfall-induced errors) were also shown to be lower in 2004 than 2006. NEXRAD- and TRMM-induced errors in near-surface soil moisture were generally of low magnitude (at the order of 0.02-0.03) and comparable to each other, while the respective CMORPH-induced error was much higher overall (at the order of 0.05-0.06), due to excessive positive bias with respect to measured soil moisture (originating from a respective bias in rainfall estimation). An inter-comparison between rainfall- and modelling-induced errors verified the results by Hossain and Anagnostou (2005a); both errors were of similar magnitude in the case of high modelling accuracy (e.g., warm season 2004), while rainfall-induced error was lower when the model's efficiency skill was relatively low (e.g., warm season 2006). Furthermore, a statistical evaluation of the scale effect on error properties revealed an increasing trend of the error propagation ratio (ratio of error in model-predicted soil moisture to the actual error in rainfall estimation) with coarser spatial resolutions.

Although the study presented useful indications about the error propagation properties of rainfall in the simulation of soil moisture, it is limited in terms of the rainfall products

investigated and LSM models. There are several satellite techniques providing high-resolution global-scale products with varying error characteristics and resolutions that need to be investigated in terms of their efficiency in the prediction of soil moisture variability. Furthermore, a point to note is that most satellite retrievals, including the two techniques presented herein, are undergoing changes aimed at improving their rain estimation error characteristics (e.g., the CMORPH technique is currently undergoing several modifications in the way morphing is performed to account for the summertime positive bias). The model used here to facilitate our investigation is only one of the several land surface schemes currently included in major land data assimilation systems. The models exhibit differences in the parameterizations devised to represent the land-atmosphere interaction processes, which would subsequently affect in a non-linear way the propagation of precipitation error in the prediction of hydrologic parameters (including the soil moisture studied herein). Those are aspects subject to future research studies.

Chapter 3 Surface and groundwater interactions in land surface modelling at the regional scale

3.1 Introduction

River-groundwater interaction at the local scale can be parameterized in terms of the difference between the river elevation and a reference groundwater top through the concept of river conductance (Rushton 2007). At the regional scale the limited resolution does not provide the sufficient information needed to apply the river conductance method. Miguez-Macho et al. (2007) generalized this approach and proposed a regional groundwater model with a parameterization of river-groundwater interactions. To circumvent the resolution problem, river conductance was parameterized as a function of the displacement of the water table depth with respect to its equilibrium value and the mean river elevation, which are computed a priori from a preliminary 1-km resolution groundwater simulation (Fan et al. 2007). However, the river-groundwater interaction itself was not validated as the study focused on addressing the sensitivity of the groundwater model with respect to the free-drainage condition. The importance of improving the river hydrology in the models is also pointed out by David et al. (2009), who integrated a vector representation of the stream and river network derived by 30 m topography in the high-resolution NOAH-distributed land surface model (Gochis and Chen 2003).

In this study, we provide observational evidences that soil moisture and surface temperature spatial distributions are related to the characteristics of the river network at regional scale (0.25 degree). At this resolution, we hypothesize that the mean effect can be modelled as a direct water flux from the river reservoir to the mean groundwater for each grid point in the same spirit of the TOPMODEL approach, but to describe the opposite process. In this framework, the sub-grid mechanisms that are responsible for the mean effect, i.e. the local infiltration at the scale of the river, the lateral hydrological processes along the river corridors, and the lateral groundwater fluxes that redistribute the local water anomaly at the grid scale, are accounted implicitly. We demonstrate the validity of this assumption by comparing observations and simulations conducted with the Community

Land Model (CLM), version 3.5 (Oleson et al. 2004; Oleson et al. 2008), that we modified to include the new parameterization.

This chapter is organized as follows: the following section (Section 3.2) includes a description of the study region and the data used. In Section 3.3 we describe the CLM model, in particular the groundwater dynamics in the original version as well as our modification that accounts for the river feedback, and the control simulation used to analyse the observed data. In Section 3.4 we discuss the observational analysis of the soil moisture and temperature data showing their dependency of the regional scale characterization of rivers. In Section 3.5 we present the results of the new parameterization introduced in CLM, and we quantify the impact of this parameterization on the water cycle and the land surface state. Discussion, conclusions and prospects for future work are provided in Section 3.6.

3.2 Study region and data

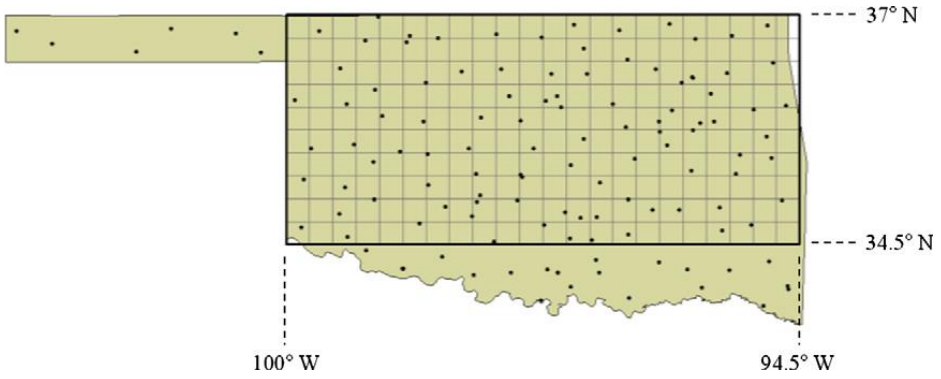


Figure 3.1. Location of the Mesonet stations in Oklahoma (black dots) and the 0.25-grid mesh over which data are interpolated and the CLM model is run.

Fig. 3.1 shows the stations locations and the mesh over which the observed data are interpolated. The study region is in the State of Oklahoma (US), where a dense network of hydrometeorological stations is present (Mesonet; Brock et al. 1995, Shafer et al. 2000). This network provides basic surface (e.g., precipitation, solar radiation and pressure) and near-surface (e.g., 10-m height temperature, relative humidity and wind) meteorological observations as well as soil moisture measurements at four different depths (5, 25, 60 and

75 cm), which comprise the focus of our analysis. A gridded version of the Mesonet database is used for the period 2000-2006, created by means of interpolating the original point meteorological observations on a 0.25-degree grid from 34.5N to 37N and 100W to 94.5W (10x22 grid). The Mesonet meteorological data were interpolated to this spatial grid using the inverse distance weighting (IDW) technique. The first four years of data (2000-2003) are used to initialize the land surface model and the last three years (2004-2006) are used for the actual soil moisture analysis and the sensitivity studies.

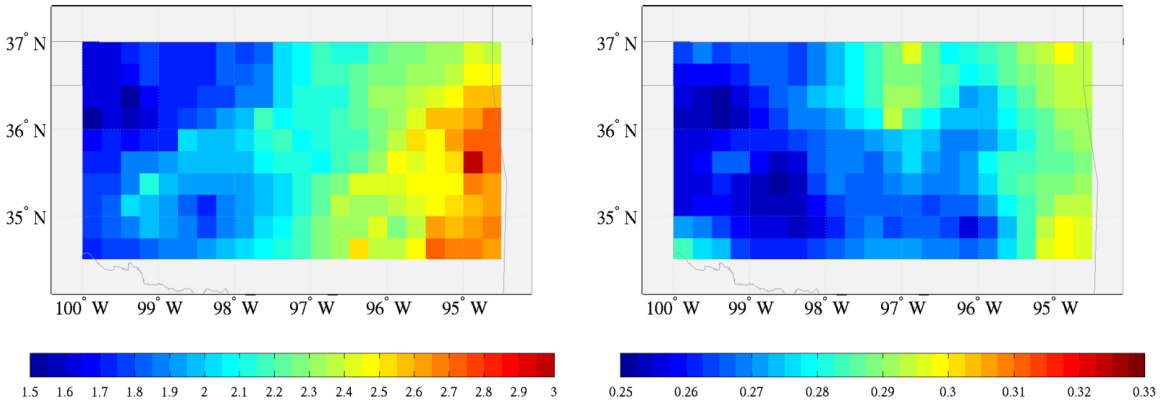
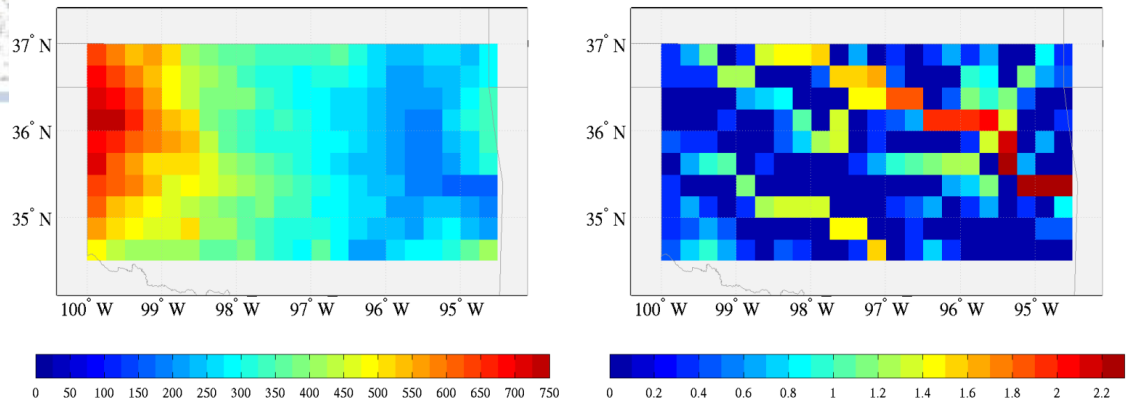


Figure 3.2. (a) Mean precipitation (mm/day) and (b) root zone volumetric soil moisture for the period 2004-2006 from the Mesonet network.

Fig. 3.2 shows the average precipitation and root zone soil moisture from January 2004 to December 2006 computed on the grid mesh presented above. The root zone soil moisture is defined as the arithmetic average of the observations at 5, 25, 60 and 75 cm depths (Teuling et al. 2006, Albergel et al. 2008). Both precipitation and soil moisture show a zonal gradient, with a drier climate in the west side of the domain and a more humid climate in the east side. Their patterns show regional features that are correlated to each other, e.g., precipitation and soil moisture exhibit local minima on the west side of the domain and local maxima in the south-east corner of the domain. The overall picture is consistent with the main forcing exerted by precipitation on soil moisture.



(a) (b)

Figure 3.3. Geomorphologic characteristics of the Oklahoma Mesonet domain: (a) Orography (m) and (b) the contributing area (see text for explanations) expressed in logarithmic scale (base 10).

Fig. 3.3a shows the average orography on the 0.25-degree grid resolution. The orography shows a main zonal gradient as well, that could be correlated to the respective gradient of the soil moisture field mentioned above. However, this soil moisture-to-orography spatial pattern dependence is overlapping with the more intense spatial patterns imposed by the precipitation forcing affecting evaporation and runoff production. Fig. 3.3b shows the common (base 10) logarithm of the *contributing area* of surface hydrology derived from the 0.25-degree topography. We define this variable as the number of grid points whose surface drainage is accumulating in the local grid point. The derivation of the contributing area depends strongly on the determination of the flow direction field. This can vary significantly depending on the algorithm and the grid resolution used. In our study, the computation of the flow direction for the 0.25-degree grid is based on the D8 single flow direction algorithm (O' Callaghan and Mark 1984). Several algorithms were tested and found capable of affecting the statistical significance of the analysis, but not the actual results. The topographic data used to calculate the flow paths, were derived by aggregating the 1/3 arc-sec National Elevation Dataset obtained from the USGS seamless database (http://seamless.usgs.gov/ned13.php). The 0.25-degree mean elevation inside the domain of our study varies smoothly between 150 and 750 m. A point to note is that at this resolution and elevation differences the maximum elevation gradient between grid points is approximately 7 m/km, which is not sufficient to generate significant lateral groundwater

fluxes on the basis of the water table head approach. We conducted preliminary tests using this approach, which showed negligible results even at 5 km grid cell resolution. As noted in Fig. 3.3b, most of the domain is covered by sub-basins of various scales discharging into the Arkansas River, which is the most prominent feature observed. The southern part of the domain is draining into the Red River. Both rivers connect to the Mississippi River, which is outside our study domain.

Land surface temperature (LST) data, based on satellite retrievals from the Moderate Resolution Imaging Spectroradiometer (MODIS), are used to provide an independent evaluation of the impact of the new CLM parameterization on the stream-groundwater interaction introduced in this study. The surface temperature data from Mesonet are not suitable for this scope, as they are part of the model forcing dataset. There is a variety of MODIS LST products available online from the Land Processes Distributed Active Archive Centre (LP DAAC; https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table), which comprises a partnership between NASA's Earth Observing System (EOS) and the United States Geological Survey (USGS). These products are created as a sequence of products beginning with a swath (scene) and progressing, through spatial and temporal transformations, to daily, eight-day and monthly global gridded products. MOD11C1.5 (Terra satellite) and MYD11C1.5 (Aqua Satellite) products (Wan 2008) were deemed suitable for our study. These are daily global LST products, which provide temperature (and emissivity) values at a 0.05-degree latitude/longitude climate model grid. They contain two LST values per day (day/night) with about 1 Kelvin accuracy, while cloud-contaminated LSTs (daytime and night-time) are removed by a double-screening method. These products also receive extensive validation through field campaigns and radiance-based validation studies (further details at http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD11).

The 2004-2006 LST data were processed to facilitate comparisons with the model LST outputs; specifically, the gridded area shown in Fig. 3.1 was cropped from the global products, and the dataset was then aggregated to the resolution used in the current study (0.25 degree). It is noted that only daytime LST values are used in the analysis, and they correspond to the times when the satellites overpass the Oklahoma area (approximately 18 UTC for Terra and 20 UTC for Aqua). The mean LST fields are shown in Fig. 3.4,

complementing the Mesonet-based 2004-2006 climatology of the domain under study discussed above. The western part of Oklahoma is characterized by a warmer and drier climate, whereas the eastern part exhibits colder and more humid soils. As noted, the soil temperature agrees with the soil moisture patterns since the soil temperature decreases for increasing water content.

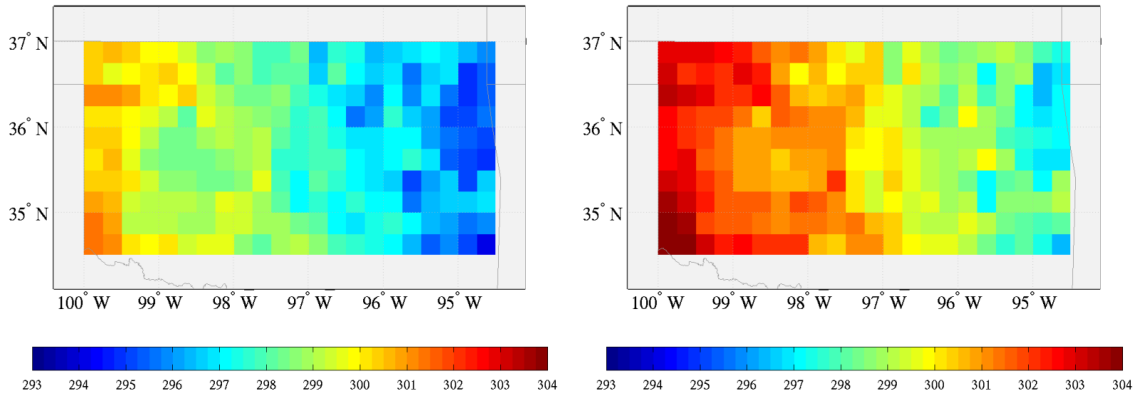


Figure 3.4. MODIS land surface temperature for the period 2004-2006 from (a) Terra satellite (18 UTC) and (b) Aqua satellite (20 UTC).

3.3 CLM formulation and the control simulation

CLM is a land surface model designed for global and regional climate applications. A description of this model can be found in Oleson et al. (2004; see also Bonan et al. 2002 for a brief history of its development), while recent developments and detailed evaluations of its performance are presented in Oleson et al. (2008) and Stöckli et al. (2008). Here, we briefly describe some details of the soil moisture and groundwater formulations. Soil moisture is resolved on a vertical grid consisting of 10 unevenly spaced layers, with the higher resolution at the surface and with the lower layer reaching the depth of 3.43m (the layer thicknesses are defined as 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33 and 113.7 cm, respectively). Soil moisture dynamics is determined by gravity and capillary force exerted between the different levels and the unconfined aquifer below the unsaturated zone (Niu et al. 2007). Water can be up-taken by plants' roots and can evaporate directly

from the upper computational layer. The water table is diagnostically defined as the separation between the unsaturated zone and the aquifer.

3.3.1 The TOPMODEL approach for runoff generation and the groundwater dynamics

Water table depth is used in the model to parameterize the sub-surface runoff production according to the TOPMODEL approach (Beven and Kirkby 1979, Niu et al. 2005). This approach is based on the assumption that subgrid spatial variability of orography exists. Therefore, for a given water table depth, ponded areas may be produced. As a consequence, the corresponding water is removed from the groundwater and is converted into runoff at a rate that decreases exponentially with the water table depth:

$$R_G = (1 - f_I) R_{G,max} exp(- d_{WT} / \Delta), \qquad \textit{(Equation 3.1)}$$

where R_G (m/s) is the runoff generation, f_I (unitless) is the fraction of ice in the soil cell, $R_{G,max}$ is the maximum runoff, d_{WT} (m) is the water table depth and Δ (m) is a decay factor related to the soil thickness. This last parameter is also involved in the computation of the maximum infiltration capacity, which affects the surface runoff production, as well as in the computation of the hydraulic conductivity of the aquifer, which modulates the interactions with the resolved soil column. In CLM (version 3.5), the parameters $R_{G,max}$ and Δ are set to 0.45 mm/s and 400 mm, respectively, while the ice fraction and the water table depth can vary during the simulation. Eq. 3.1 states that the maximum runoff is produced when the water table approaches to the surface. It decreases with increasing water table depth, and eventually converges to zero if the water table is very low, which is consistent with the case that all subgrid orography is much higher than the water table and there are no ponded areas. The parameters used in Eq. 3.1 are constant across the entire model domain. CLM accounts for the spatial variability of hydraulic conductivity of the soil layers and the aquifer through the dependency of these parameters on the soil texture.

The temporal variation of the groundwater (W_A) stored in the unconfined aquifer is defined as:

$$dW_A/dt = Q - R_G \qquad \textit{(Equation 3.2)}$$

where Q (m/s) is the recharge of water from the unsaturated soil (Niu et al. 2007). For a constant recharge rate and unfrozen soil ($Q = Q_0 = const$ and $f_I = 0$), (2) and (1), after imposing also $dW_A/dt = 0$, predict an equilibrium water depth ($d_{WT,eq}$) that is given by:

$$d_{WT,eq} = \Delta \ln (R_{G,max}/Q_0) \qquad \textit{(Equation 3.3)}$$

This formulation is local, i.e. the lateral groundwater fluxes between model grid-points are neglected. Surface runoff in CLM is then transported by a river model (Branstetter 2001), which is implemented on an independent grid communicating with the CLM grid trough the Model Coupling Toolkit (MCT; Larson et al. 2005, Jacob et al. 2005) interpolation library. The fact that CLM includes rivers and groundwater schemes facilitates the implementation of the parameterizations described in the following sections, which is one of the main reasons that the particular land model was chosen in this study. The extensive model development and validation in Oklahoma region that we have conducted in previous studies (Serpetzoglou et al. 2010, Wang et al. 2005; 2006; 2009a) further justifies our choice.

3.3.2 The parameterization of river-groundwater interaction

The groundwater equation (Eq. 3.2) is modified to account for the recharge from the river water (Q_R):

$$dW_A/dt = Q - R_G + Q_R \qquad \textit{(Equation 3.4)}$$

On the basis of this modification, the equilibrium water table depth equation (Eq. 3.3) becomes:

$$d_{WT,eq} = \Delta \ln (R_{G,max}/ (Q_0 + Q_{R0})), \qquad \textit{(Equation 3.5)}$$

where it is also assumed constant recharge rate from rivers ($Q_R = Q_{R0} = const$).

Our consideration, supported by the data analysis presented in the following section, is that increase of soil moisture is positively correlated with the contributing area of surface hydrology, which can be represented by the simulated amount of river water per grid cell. Therefore, we propose a parameterization of this additional groundwater recharge term as a function of the river water volume:

$$Q_R = a\, W_R^b, \qquad \textit{(Equation 3.6)}$$

where W_R is the total river water volume in the grid cell as computed by the river transport model included in CLM. Eq. 3.6 is very general and states than the recharge from rivers increases with the amount of river water, which is consistent with the common sense. The validity of this equation and specific values of the involved parameters will be discussed in Section 3.5, where we compare the results of CLM with the new parameterization against the observations.

3.3.3 CLM implementation and control simulation

We implemented CLM and RTM in the grid shown in Fig. 3.1. Simulations (CLM-CTRL) were initiated on January 1, 2000 and lasted seven years (2000-2006). Similarly to Serpetzoglou et al. (2010), the first four years were used for the model spin up and are not accounted in the analysis. We used the Mesonet meteorological data (see section 3.2 for details) interpolated on the same grid to provide the surface boundary conditions for the model. CLM requires vegetation, soil type and soil texture data that are provided at high resolution (0.05 degree), based on recently developed MODIS products (Lawrence and Chase 2007). In fact, CLM includes a hierarchical structure of variables that accounts for the subgrid distribution of certain parameters as land cover.

The average fields of root zone soil moisture, water table depth, land surface temperature and surface energy fluxes computed on the last three years of the CLM control simulation (CLM-CTRL) for period 2004-2006 are presented in Fig. 3.5. For the derivation of root zone soil moisture from the multilayer model subsurface estimates, a weighted vertical interpolation is applied over the model layers that are within the top 1 m of the soil column (following Kumar et al. 2009). The weights are extracted based on the layer

thicknesses, thus taking into account the first eight model layers (the 8th model layer output participates in the weighted average only to a degree proportional to the part of the layer that is above 1 m depth).

The simulated root zone soil moisture field (Fig. 3.5a) can be directly compared with the corresponding field from the observations (Fig. 3.2b). The model exhibits an overall bias that could depend on the means of calculating the root zone soil moisture values from the multilayer subsurface observations or model outputs. For instance, the bias values would change if the observed root zone soil moisture was determined as the soil moisture in the 75 cm depth (graph not shown here), instead of the arithmetic average of the four Mesonet depths shown in Fig. 3.2b. The various techniques of estimating the root zone soil moisture as well as the observed data-model bias are not the focus of this study and could be evaluated in future work. On the other hand, the temporal variability of the computed near-surface soil moisture (shown in Serpetzoglou et al. 2010) is well correlated to the observed soil moisture values in terms of domain averages.

The water table depth field (Fig. 3.5b) shows a highly variable spatial pattern, which seems to be consistent overall with the respective pattern of both observed and simulated root zone soil moisture. The simulated LSTs at 18 and 20 UTC (Figs. 3.5c and 3.5d) exhibit good correspondence with the respective MODIS LST fields shown in Figs. 3.4a and 3.4b. Specifically, we note low biases and high correlation with respect to the zonal gradients. As an estimate of LST we used the temperature of the upper most soil layer. The results that we present do not change significantly if we use the canopy temperature instead (see Wang et al. 2009b for a more elaborate discussion on this issue). The average fields of surface sensible (Fig. 3.5e) and latent heat flux (Fig. 3.5f) are also characterized by intense zonal gradients consistently with the distribution of their primary modulators, i.e., surface temperature and soil moisture.

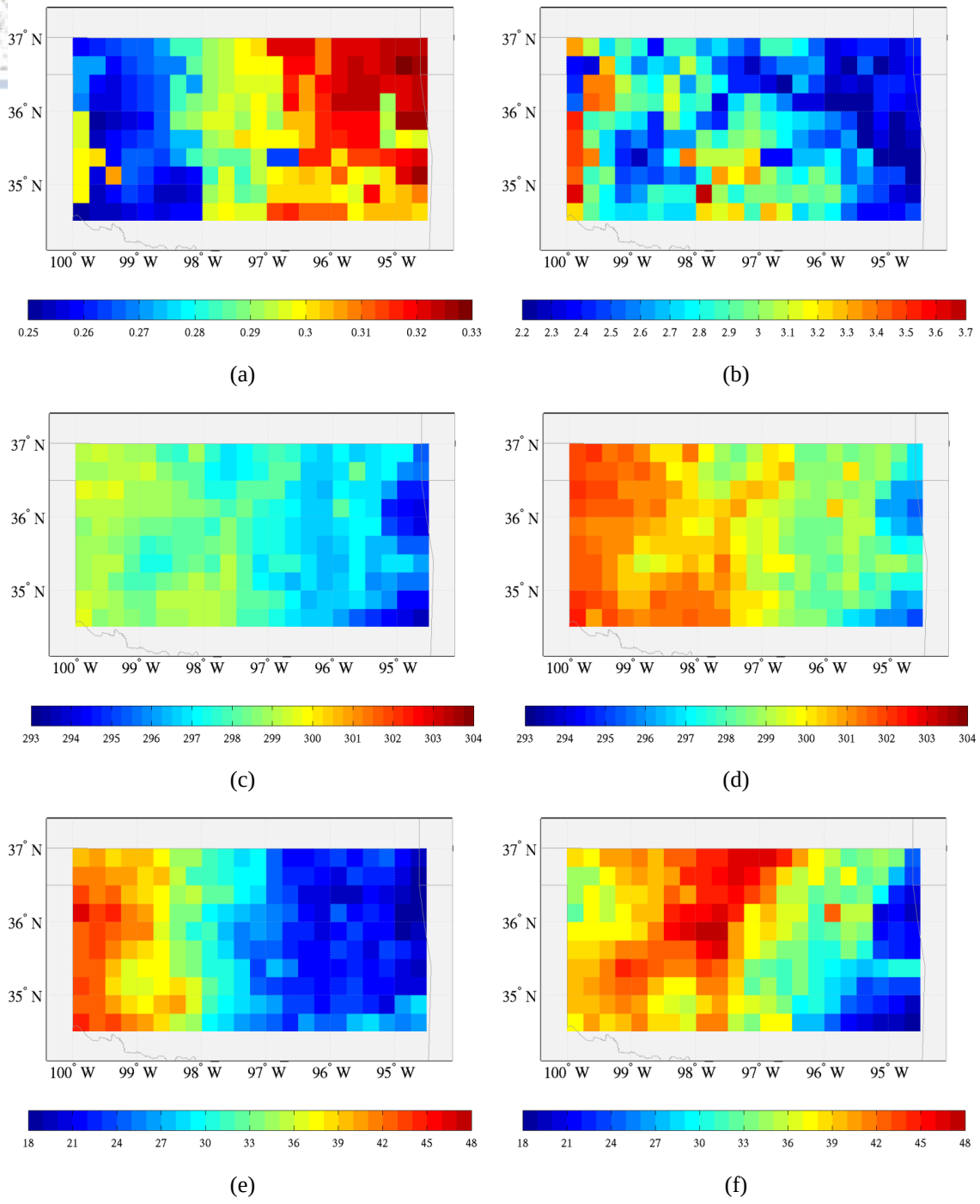


Figure 3.5. Average fields for the period 2004-2006 from the CLM-CTRL simulations. (a) Root zone soil moisture (%v/%v), (b) water table depth (m), (c) ground temperature at 18 UTC (K), (d) ground temperature at 20 UTC (K), (e) sensible heat flux (W/m^2), and (f) latent heat flux (W/m^2).

3.4 Soil Moisture and Surface Temperature Analysis

In order to evaluate the effect of rivers on the spatial patterns of soil moisture and surface temperature based on measured data, we need to filter out the spatial patterns imposed by the meteorological forcings and modulated by vegetation and soil characteristics. This is accomplished by normalizing the observations with the simulated fields obtained by the original CLM. In fact, CLM is able to reproduce the effects of precipitation, vegetation and soil characteristics on soil moisture and temperature patterns, while it does not account for the groundwater recharge from river water. Normalization of the observed soil moisture with the simulated values from CLM-CTRL is applied directly on the Mesonet data as spatial interpolation of the data would smooth out the spatial patterns that we seek to analyse, especially at deeper levels where data are sparser. Instead, we interpolated the CLM-CTRL values at each station location using bilinear interpolation on the model data. On the other hand, the land surface temperature analysis is conducted over all the grid points in the domain, due to the more comprehensive spatial coverage of the satellite observations. In both cases, we selected only the points that are well correlated with the model results to obtain a meaningful normalization. The threshold for correlation was set to 0.65, which was found to be a good compromise of maintaining an adequate size of the sample data, critical especially for the soil moisture data. The analyses are restricted to periods when the soil is not frozen to avoid the masking effect on soil moisture due to the reduced infiltration (Cherkauer and Lettenmaier 1999).

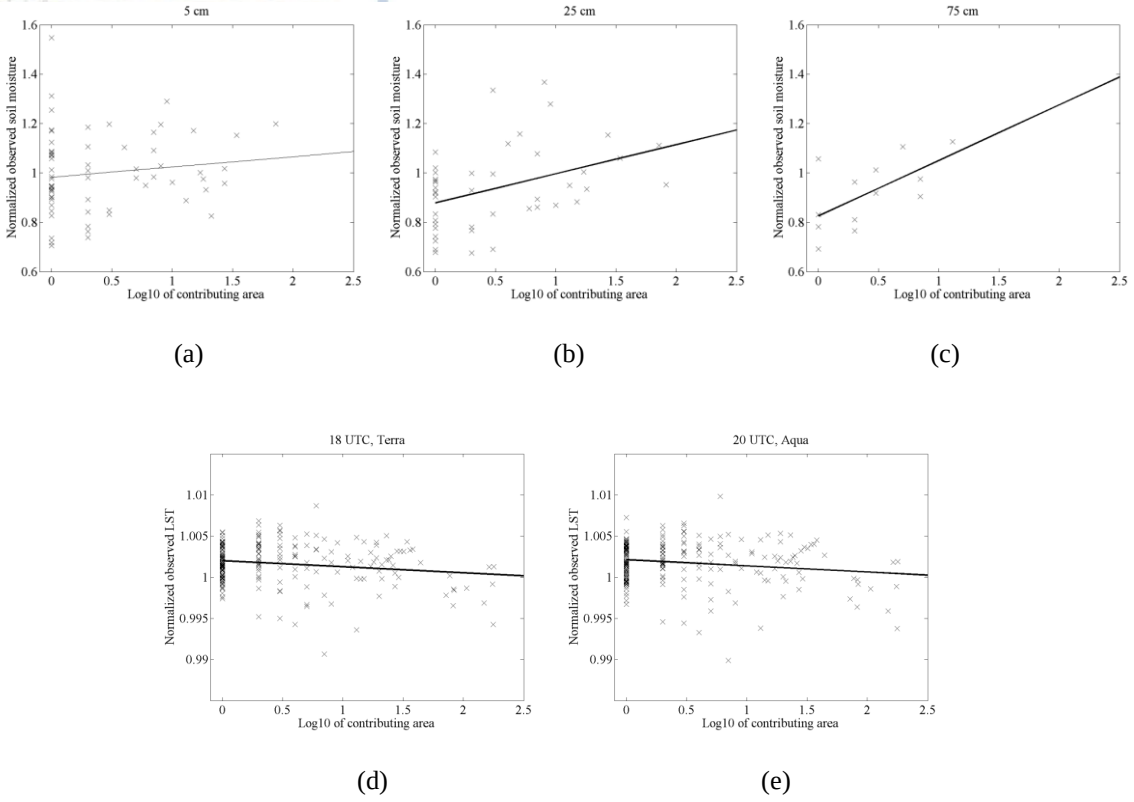


Figure 3.6. Observed soil moisture on the Mesonet stations normalized with the simulated CTRL soil moisture (averages over the period 2004-2006) vs. the common logarithm of the contributing area (as shown in Fig. 3.3b) for (a) 5 cm, (b) 25 cm and (c) 75 cm. Observed LST normalized with the simulated CTRL LST (averages over the period 2004-2006) vs. the common logarithm of the contributing area for (d) Terra overpass (18 UTC) and (e) Aqua overpass (20 UTC). Straight lines visualize the corresponding linear regression. Bold lines are used where the regression is statistically significant. Table 3.1 shows the detailed results of the regression analysis.

The results of this procedure are presented in Fig. 3.6, where the normalized soil moisture and surface temperature are plotted against the contributing area shown in Fig 3.3b, which is representative of the river distribution. Besides a large scatter in the normalized data, these plots suggest that the river distribution actually plays a role in the spatial organization of soil moisture and surface temperature at the regional scale. In fact, the normalized soil moisture tends to increase with the contributing area, while surface temperature decreases. This means that the model exhibits a dry and warm bias with respect to the observations in the grid points corresponding to the rivers. The statistical significance of this results is computed after the regression analysis that is summarized in Table 3.1.

Table 3.1. Results of the regression analysis conducted on the normalized soil moisture and surface temperature versus the common logarithm of the contributing area (data shown in Fig. 3.6). Normalization of observed data is performed with the corresponding fields obtained by the CTRL model simulation. For each depth or overpass the corresponding values of the intercept minus one (I-1) and the slope (S) are given, followed by the respective standard deviations and the thresholds of the statistical significance.

	I - 1	**Sigma I.**	**Signif. I.**	**S**	**Sigma S.**	**Signif. S.**
05 cm SM	-0.017	0.025	< 90 %	0.042	0.037	< 90 %
25 cm SM	-0.121	0.031	> 99 %	0.118	0.041	> 99 %
75 cm SM	-0.174	0.048	> 99 %	0.225	0.087	> 95 %
Terra LST	0.0020	0.0002	> 99.9 %	-0.00073	0.00028	> 99 %
Aqua LST	0.0021	0.0002	> 99.9 %	-0.00075	0.00030	> 95 %

In Table 3.1, the intercepts of the computed regression are listed with a unit subtracted; meaning that a perfect match for the normalized value (1) would be achieved if zero is approached. These values are consistent with the mean wet and cold bias of CLM respect to the observations, which is not the focus of the present study. We use the slopes resulting from the regression analysis to quantify the relationships between the coarse scale characterization of the basin and the soil moisture and temperature values. For soil moisture, at two out of the three depths presented a statistically significant positive slope is found, meaning that the normalized soil moisture tends to increase as a function of the contributing area. In particular, there is less than 1% probability to have obtained the result at 25 cm depth by chance, while the statistical significance at the 75 cm is lower. The relationship between soil moisture and contributing area is not statistically significant at the surface layer (5 cm depth), despite the larger number of available data, which could be attributed to the fact that the upper layer soil moisture is mainly affected by surface fluxes (precipitation and evaporation). The results in terms of surface temperature are even more solid due to the higher amount of points included in the analysis. A statistically significant relationship of 99% is obtained for the Terra overpass at 18UTC. The significance decreases, but results still hold, for the Aqua overpass at 20UTC. The computed trends are negative, which means that the original CLM model tends to predict warmer surface temperatures than the satellite observations in the grid points corresponding to rivers. This effect is closely linked to the soil moisture pattern. In fact, during daytime, solar radiation

warms up the soil that emits back in the infrared and generates sensible and latent heat fluxes. Higher soil moisture values are known to enhance evapotranspiration and decrease surface temperature, especially during daytime. This explains the model overestimation of surface temperatures in grid cells with high contributing area (negative slope in Table 3.1) and why this relationship is less significant in the evening (20UTC) when the solar forcing reduces.

It is noted here that we were not able to obtain any significant relationship using other geomorphologic variables such as the terrain elevation or the Laplacian of the terrain elevation to diagnose lateral groundwater fluxes among different grid cells. The slope at 75 cm is suspiciously high and it would be unphysical to attribute it only to the effect of rivers. However, it suggests that soil moisture dependency of contributing area increases with depth. Therefore, this effect is presumably linked to the groundwater dynamics. Moreover, we verified that the statistically significant relationship presented above is not artificially introduced by the model. In fact, the simulated soil moisture does not show any significant (negative) slope with respect to the contributing area that could influence our analysis (see Fig. 3.7). In this figure, soil moisture at 25 and 75 cm shows high variability and no trend for low to moderate values of the contributing areas. Actually, soil moisture increases at the high contributing area cells because of the higher precipitation in the eastern part of the region (see Fig. 3.2a). Nevertheless, comparison with the observations shows that in some grid cells even higher values should be expected because of the interaction with the river water.

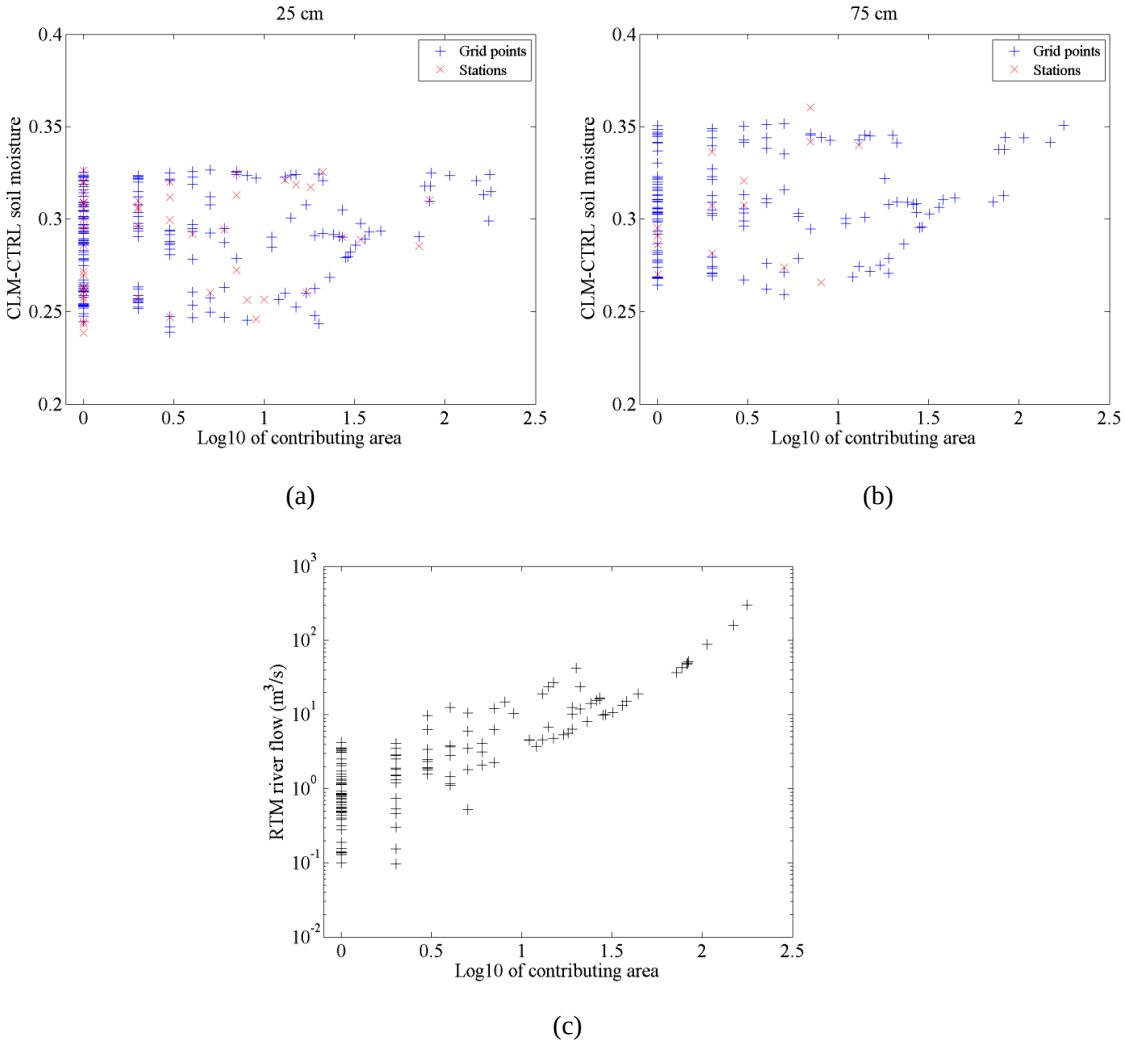


Figure 3.7. Top panels: CLM-CTRL volumetric soil moisture at (a) 25 and (b) 75 cm depths vs. the common logarithm of the contributing area. Values are shown for all points of the domain (blue) and interpolated on the station locations (red). Bottom panel: RTM river flow in all grid points vs. the common logarithm of the contributing area. All values are obtained by averaging over the period 2004-2006.

The hydrodynamic variable that is more closely related to the contributing area is the amount of river water. In nature, rivers are usually directly connected to groundwater below the river bed. Horizontal groundwater fluxes redistribute laterally the groundwater anomalies generated at the scale of the river network. Thus, water transported by rivers can affect groundwater, and consequently soil moisture, over an area that could be quite larger than the area occupied by the river. Arguably, this process is responsible for the signals that we identified in the data. In fact, we have shown that the model bias with

respect to the observations is closely related to distribution of rivers (i.e. the contributing area) over areas of more than 500 km^2, which is the order of magnitude of the grid cell size used in our analysis. This is the effect we will model in CLM with the new parameterization introduced in the previous section. A point to note is that the current analysis is based on the long-term averages of soil moisture; therefore, at this point, it does not provide information on the temporal variability of this process, but it accounts only for the spatial distribution of the resulting soil moisture and surface temperature regional patterns. In the next section, we apply the new CLM parameterization aimed to reproduce the prescribed results and assess its impact on representing the climatic patterns of soil moisture and surface temperature spatial variability.

3.5 Results of the modified CLM

3.5.1 Fine tuning and application of the new parameterization

The results discussed in the previous section strongly suggest that infiltration of river water into groundwater affects the regional scale soil moisture field and the surface temperature patterns. We propose to account for this phenomenon by applying in the current CLM configuration Eq. (3.4) and (3.6), which involves the total amount of surface water as basic variable. Specifically, Eq. 3.4 represents in a general form the dependency of the flux of surface water on the total amount, disregarding other possible factors. Fig. 3.7c shows the mean river flow from the river transport model coupled to CLM. The values are shown against the contributing area of the corresponding grid cell. The dependency of river water volume on the contributing area shown in Fig. 3.7c suggests that the proposed modification - using a positive coefficient - would likely produce a positive correlation between soil moisture and contributing area. We explore two approaches for the specific functional form of Eq. 3.6, namely, two values for the exponent b are tested: b=1 and b=2/3, corresponding dimensionally to the parameterization of the recharge term as a function of the *volume* of river water and as a function of the *surface* of river water, respectively. Depending on the exponent value, the coefficient “a” would have dimensions of an inverse timescale or conductivity, respectively. In the following, we explore the

effect of these parameterizations in terms of soil moisture, which is more closely related to the groundwater dynamics, providing at the same time a basis for fine tuning the proposed modifications for reproducing the results of surface temperature that accounts for a larger portion of the domain. The proposed parameterizations are expected to eliminate the dependence observed in Fig. 3.6; the observed soil moisture and surface temperature data for the various cases under study normalized with the results obtained by the modified model should be characterized by no slopes (or statistically insignificant slopes at least). An alternative way to obtain the same result would be to show that the slopes produced by normalizing soil moisture and surface temperature values from CLM with the new parameterization to the control CLM simulations match the slopes demonstrated in Fig. 3.6 based on observed values. This evaluation will form the basis to argue that the new parameterization can account for the previously neglected, yet significant, effect of surface drainage to groundwater.

We subjectively evaluated the model parameters by iterating for a range of values and comparing the model derived climatology against the observed spatial soil moisture patterns. For the first parameterization (b=1) the parameter value for "a" is 10^{-9} (s^{-1}). Since this parameter is dimensionally an inverse time scale, it can be interpreted as the time scale necessary for the groundwater anomaly produced at the river scale to affect soil moisture in an area of 0.25 degree by 0.25 degree. It corresponds roughly to a time of 30 years, suggesting a very slow process. However, this is not a precise estimate because it is based on a closure parameter in a highly parameterized model; nevertheless, it implies that the time scale of the involved process could be of the order of several years. This process is identifiable in the observations because rivers are persistent features and do not change much over this time scale.

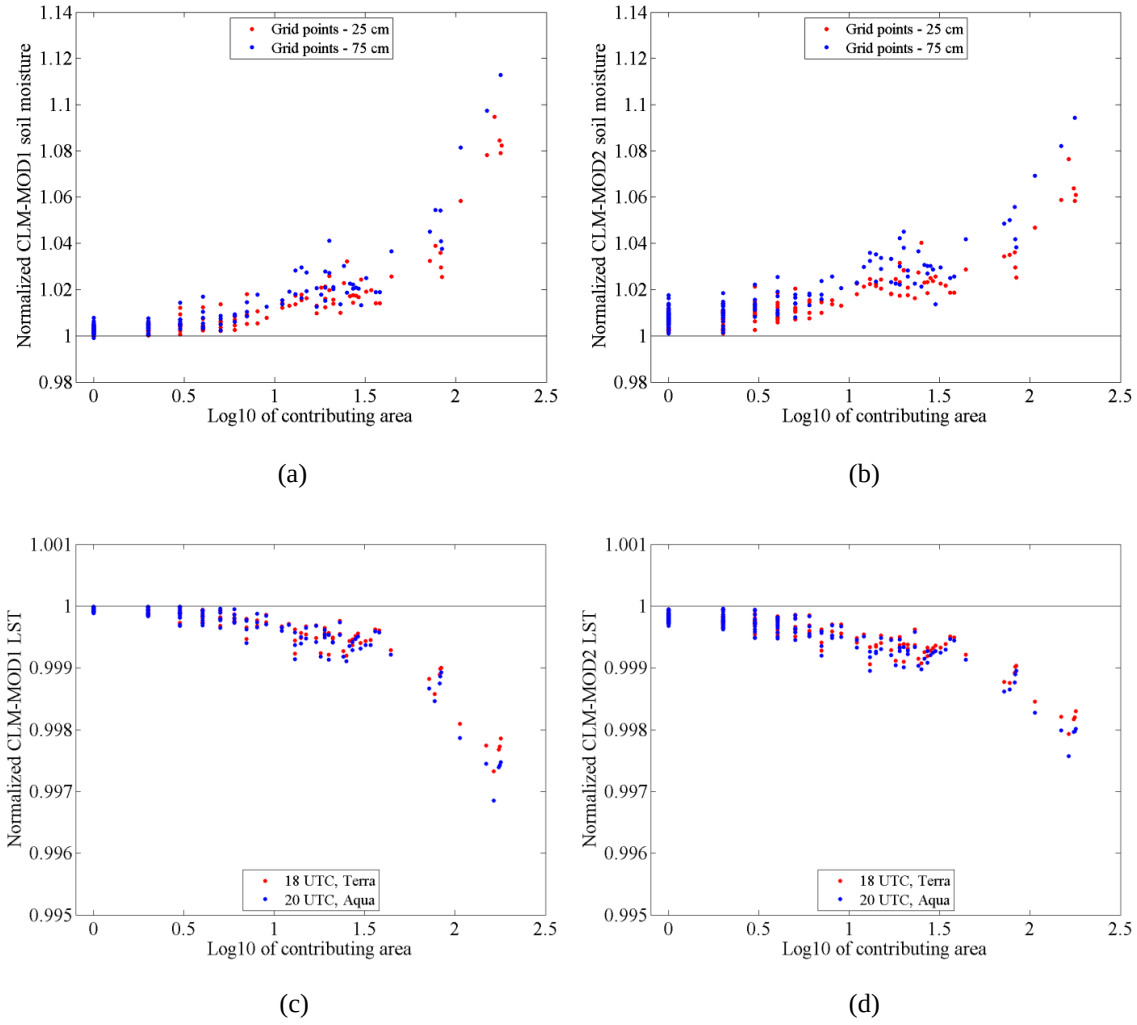


Figure 3.8. Upper panels: Modified model soil moisture for two depths (25 cm and 75 cm) normalized with the respective CTRL simulation values vs. the common logarithm of the contributing area for (a) the stream-groundwater flux defined proportional to the river water volume ($a=10^{-9}$; CLM-MOD1) and (b) the stream-groundwater flux defined proportional to the surface of river water ($a=2\cdot10^{-7}$; CLM-MOD2). Bottom panels: As in upper panels but showing the normalized surface temperature outputs at 18 UTC (Terra overpass) and 20 UTC (Aqua overpass) from (c) CLM-MOD1 and (d) CLM-MOD2. All values represent 2004-2006 averages.

Fig. 3.8a shows the results of CLM with the new parameterization of b=1 (the respective simulations will be hereafter referred to as CLM-MOD1) for two different depths. CLM-MOD1 is effective in producing qualitatively the desired statistics. Specifically, the induced anomalies are positively correlated to the contributing area as we found in the observations, while the values for slopes are of the same order with those shown in Fig 3.6. Moreover, the comparison between different depths shows that the effect

is more pronounced at 75 cm than at 25 cm, which was also suggested by our observational analysis. The new parameterization does not correct the overall positive bias exhibited in the CLM-Mesonet comparisons, in fact, no attempts are made here towards correcting the intercept of the regression analysis to a value close to 1. This could be achieved by modifying other parameters of CLM, but this is beyond the scope of our study.

Fig. 3.8b shows the results of the parameterization of the stream-groundwater flux as a function of the surface of river water (b=2/3). These simulations will be hereafter referred to as CLM-MOD2. The closure parameter is set to $a=2\cdot10^{-7}$ (m/s) producing a similar effect to the parameterization discussed in the previous paragraph. The numerical value of this parameter cannot be related physically to the actual conductivity at the scale of the river: according to our conceptual framework, this value is the conductivity that would produce the same overall effect in an idealized situation in which the river water is uniformly distributed over the gridded area. This parameterization produces more linearly distributed values, thus better similarities with the observations. However, the scatter associated with the observed data does not allow us to determine which of the functional dependencies better fits the data. Differences between the two parameterizations are small compared to the impact introduced in terms of the spatial pattern of soil moisture. Thus, the new parameterization, with any of the two proposed functional forms, can reproduce the observed process with an acceptable accuracy, without the need to consider any information from the sub-grid topography and the actual river distribution. This is one of the main outcomes of our study.

The respective comparisons of normalized surface temperature outputs at 18 UTC (e.g., Terra overpass) and 20 UTC (e.g., Aqua overpass) produced by CLM-MOD1 and CLM-MOD2 simulations are shown in Figs. 3.8c and 3.8d, respectively. As in soil moisture, the new parameterization does not correct the mean bias of normalized surface temperature, but it reproduces the slope characterizing the normalized observations as depicted versus the contributing area in Figs. 3.6d and 3.6e as expected. In fact, both the proposed parameterizations show a similar behaviour compared to the observations. CLM-MOD2 normalized surface temperature is more linearly distributed over the span of the contributing area values and exhibits less variations at the high contributing area cells compared to CLM-MOD1.

The previous discussion is strengthened when depicting the observed soil moisture and surface temperature data in the same way as shown in Fig. 3.6, but in this case observed data are normalized against the values simulated by CLM with the new parameterization where b=2/3 (CLM-MOD2). The results are graphically presented in Fig. 3.9 and the respective regression analysis is summarized in Table 3.2. As mentioned previously, the slopes of the linear regression should tend to zero as the effect that caused them is now accounted in the new parameterization. In fact, the herein slopes, with the exception of the one for the 75-cm soil moisture, are not statistically significant. This indicates that the new parameterization is actually able to reproduce the observed effect, especially for the cases of 25-cm deep soil moisture and surface temperature at 18UTC (Terra overpass), which comprised the most significant discrepancies between observations and the original model formulation based on the results of Fig. 3.6 and Table 3.1. It is not surprising that the slope at 75 cm persists even with the new parameterization; as we already noted, this strong gradient at 75-cm depth cannot be completely attributed to the river effect. A similar result is obtained by normalizing the observed data with the CLM-MOD1 experiment (not shown).

Table 3.2. As in Table 1, but for the data shown in Fig. 10 (normalization performed with the CLM-MOD2 model simulation values).

	I-1	**Sigma I.**	**Signif. I.**	**S**	**Sigma S.**	**Signif. S.**
05 cm SM	-0.027	0.026	< 90 %	0.019	0.037	< 90 %
25 cm SM	-0.098	0.033	> 99 %	0.051	0.041	< 90 %
75 cm SM	-0.171	0.050	> 99 %	0.220	0.087	> 95 %
Terra LST	0.0021	0.0002	> 99.9 %	-0.00024	0.00028	< 90 %
Aqua LST	0.0022	0.0002	> 99.9 %	-0.00020	0.00029	< 90 %

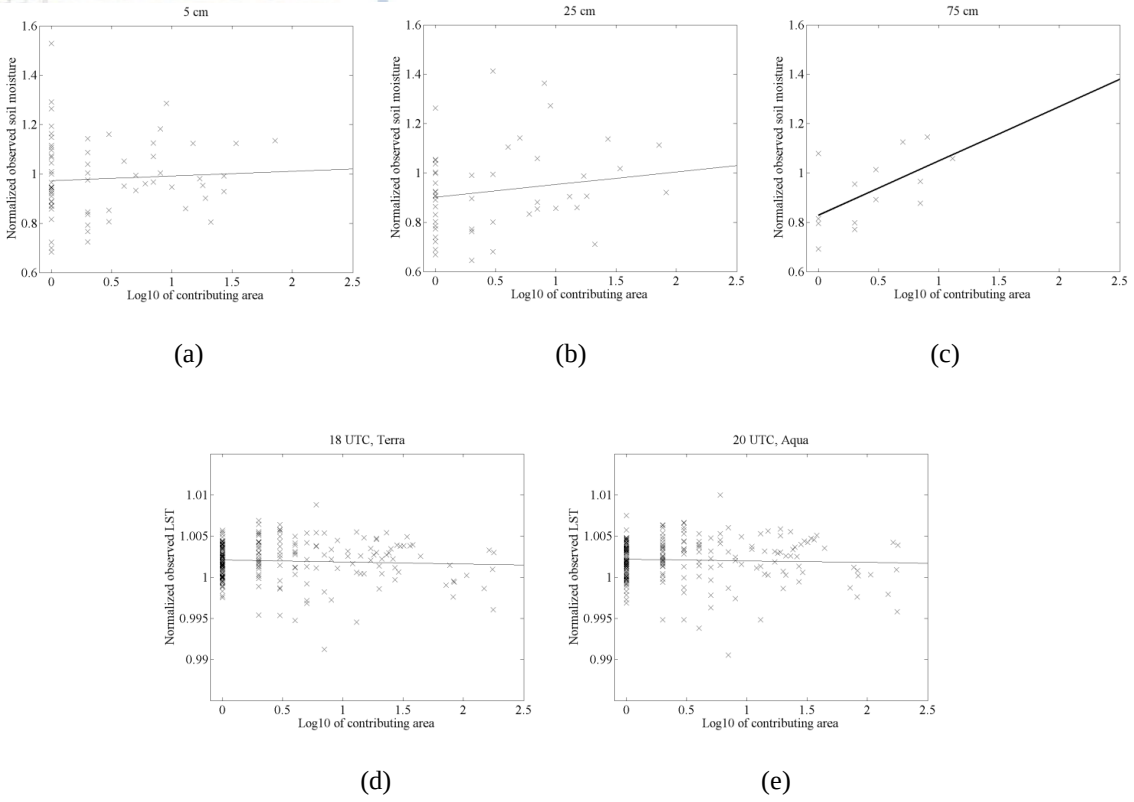


Figure 3.9. As in **Figure 3.6**, but with the respective data normalized with the simulated CLM-MOD2 soil moisture and surface temperature. **Table 3.2** shows the detailed results of the regression analysis.

It is worth mentioning that several experiments were conducted with alternative functional forms regarding the prescribed parameterization. For instance, one experiment was to include the grid cell conductance, or the water table depth, to resemble the river conductance method, which is known to describe more accurately the physics of ponded infiltration occurring at the river scale. However, at the spatial resolution used in our study the results do not improve or are even degraded. Arguably, at the regional scale the dynamics are different. A physical interpretation of the regional scale processes is difficult to obtain due to the number of unresolved hydrological processes that may be involved in the different climate and weather regimes accounted in a bulk way. The relationships presented herein use statistical results from dimensional analysis to represent the amount of river water needed to reproduce the observed data. This is an accepted method to use when the underlying physics is unclear (Buckingham 1914).

3.5.2 Impacts on water cycle and land surface state

In this section we quantify the effects of the new parameterization on the water cycle and the land surface state. The two proposals exhibit qualitatively similar results. Here we show only the results of CLM-MOD2, for brevity. The differences on root zone soil moisture, water table depth, and surface energy fluxes produced by the new parameterization with respect to the original CLM for the period 2004-2006 are presented in Fig. 3.10. The newly developed parameterization increases the spatial variability of all aforementioned variables. The induced anomaly is more intense for grid cells with high contributing area (e.g., the river network), as expected. Grid cells that are far away from streams are characterized by the least induced variability (close to zero). A quantitative analysis of the observed effect with respect to high contributing area cells shows that the effect by the new parameterization reaches the order of 1 m for water table depth (35% difference), 3% for volumetric root zone soil moisture (10% difference), and 5 W/m^2 for sensible and latent heat fluxes (25% and 15% difference, respectively).

A comparison between the Terra and Aqua LST data with the CLM-CTRL and CLM-MOD2 simulations is summarized in Fig. 3.11. From these figures it is evident that the control CLM overestimates surface temperatures in areas including, and surrounding, the river network, while it underestimates surface temperatures elsewhere. Positive and negative biases reach the values of ±2 deg K, respectively. The effect of the proposed CLM parameterization on surface temperature is presented in Figs. 3.11c and 3.11d, with regard to the Terra and the Aqua overpass times, respectively. As noted, the modified CLM produces colder temperatures up to 0.7 degree on the river routes, thus reducing the CTRL-induced bias by up to 35%.

So far we have based our analysis on 3-year averages of all variables taken into account, observed data and model outputs. In an attempt to evaluate the year-to-year variability of the impact induced by the proposed CLM parameterization, the analysis described above was repeated in terms of root zone soil moisture and land surface temperature model outputs, with the exception that separate averages were computed for each year. The results are presented in Fig. 3.12 for the years 2005 and 2006. Indeed, a significant difference is observed between the respective results for 2005 and 2006, both

for soil moisture and surface temperature. For 2005, the modified CLM produces up to 3.6% wetter root zones and up to 1.1-deg K colder temperatures on the high contributing area cells as compared to the CTRL simulations, whereas for 2006, these differences reach only up to 0.7% and 0.3 deg K, respectively. Similar variability is exhibited for the other variables (e.g., water table depth, surface energy fluxes) examined earlier in the study (graphs not shown here). The respective analysis for 2004 (graphs not shown here) reveals even better modified CLM-induced improvement than 2005 for root zone soil moisture (up to 4% wetter soil column on river routes), yet lower differences than 2005 with respect to land surface temperature (up to 0.6 deg K colder temperatures on the river network).

The prescribed results could be mainly attributed to the climatological wetness difference between the three years under consideration (see also chapter 2 for summer season comparisons); the 2005 relatively dry climate succeeded a wet year 2004 (about 35% more precipitation on average with respect to 2005 and 2006), that was followed by a similarly dry 2006. In fact, the wet conditions that prevailed throughout 2004 and especially during the post-summer period resulted in sustaining high soil moisture values during most of 2005 despite its reduced precipitation overall, while this lag was discontinued in 2006. Therefore, it is inferred that the proposed CLM parameterization produces the stronger effect during wet and wet-preceded climates, whereas its efficiency is reduced during dry conditions. These results can be considered directly equivalent with the outcomes of the error propagation study analysed in chapter 2 regarding the overall CLM performance in predicting soil moisture variability; as seen in paragraph 2.4.2, CLM had performed better during the wet summer 2004 as opposed to the dry summer 2006. These features are important to climate modelling, as they affect the efficiency of LSM on quantifying the land surface states and water cycle variability and could be further investigated in a future study.

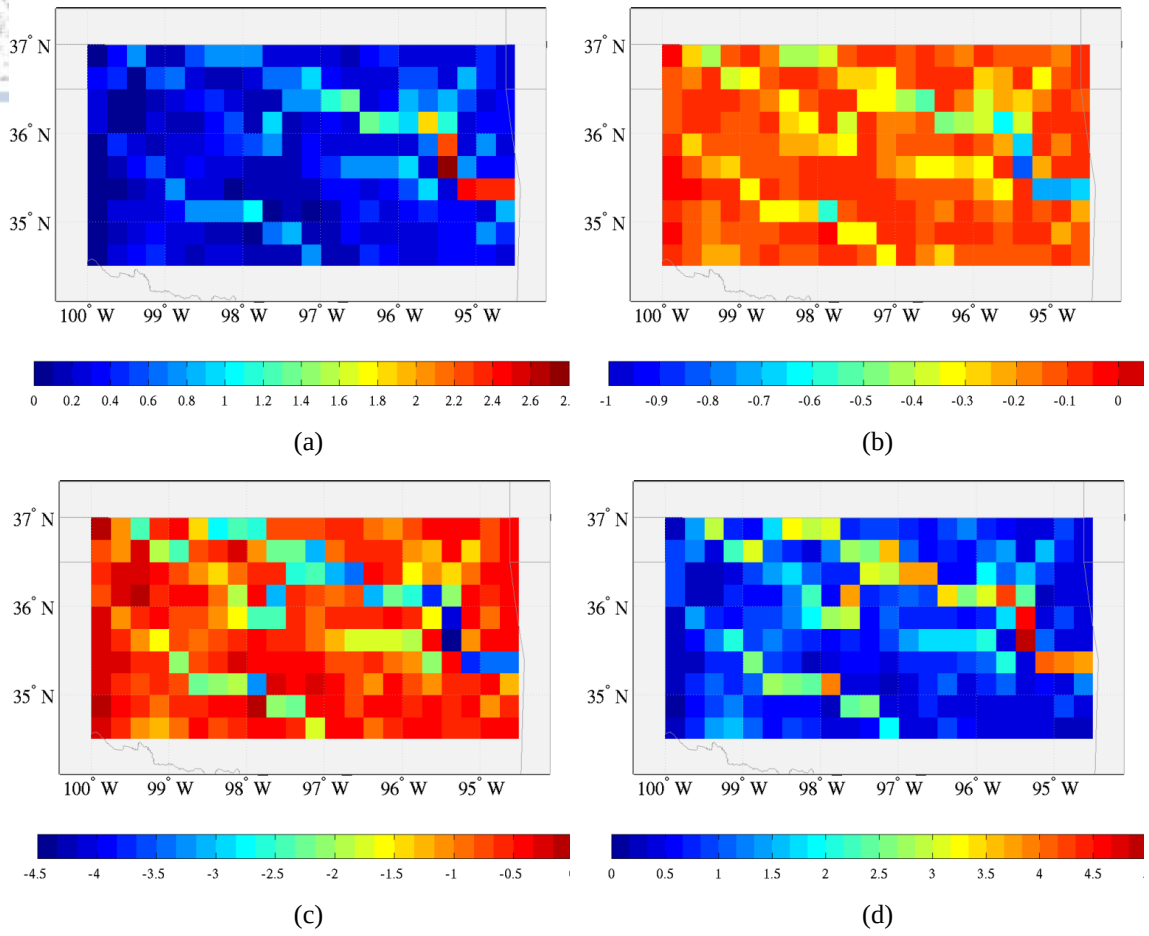


Figure 3.10. Sensitivity to the new parameterization of (a) volumetric root zone soil moisture (%), (b) water table depth (m), (c) sensible heat flux (W/m^2) and (d) latent heat flux (W/m^2). Panels show the respective differences between CLM-MOD2 and CLM-CTRL simulations.

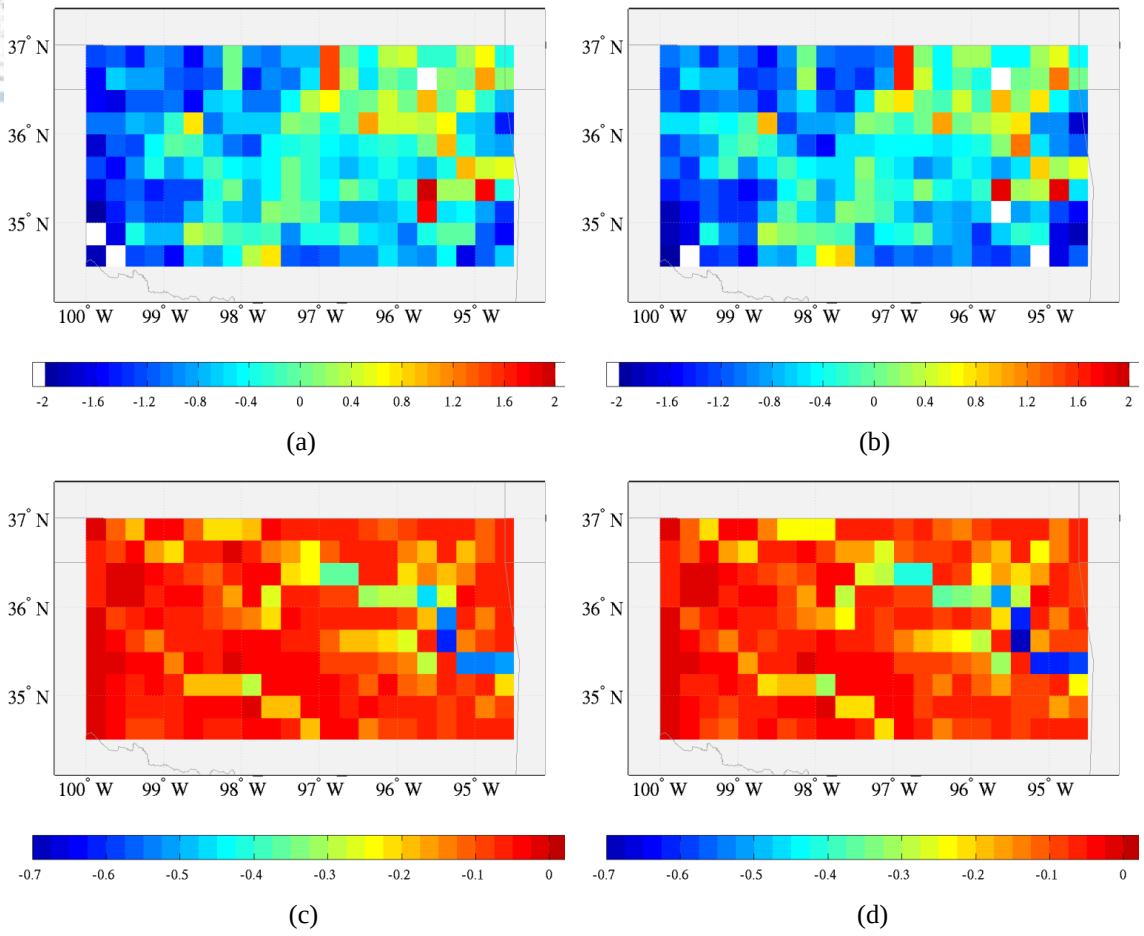


Figure 3.11. Upper panels: Differences in mean LSTs for 2004-2006 between (a) CLM-CTRL run and Terra observations (18 UTC) and (b) CLM-CTRL run and Aqua observations (20 UTC). Lower panels: Differences in mean LSTs for 2004-2006 between CLM-MOD2 run and CLM-CTRL run at (c) 18 UTC and (d) 20 UTC.

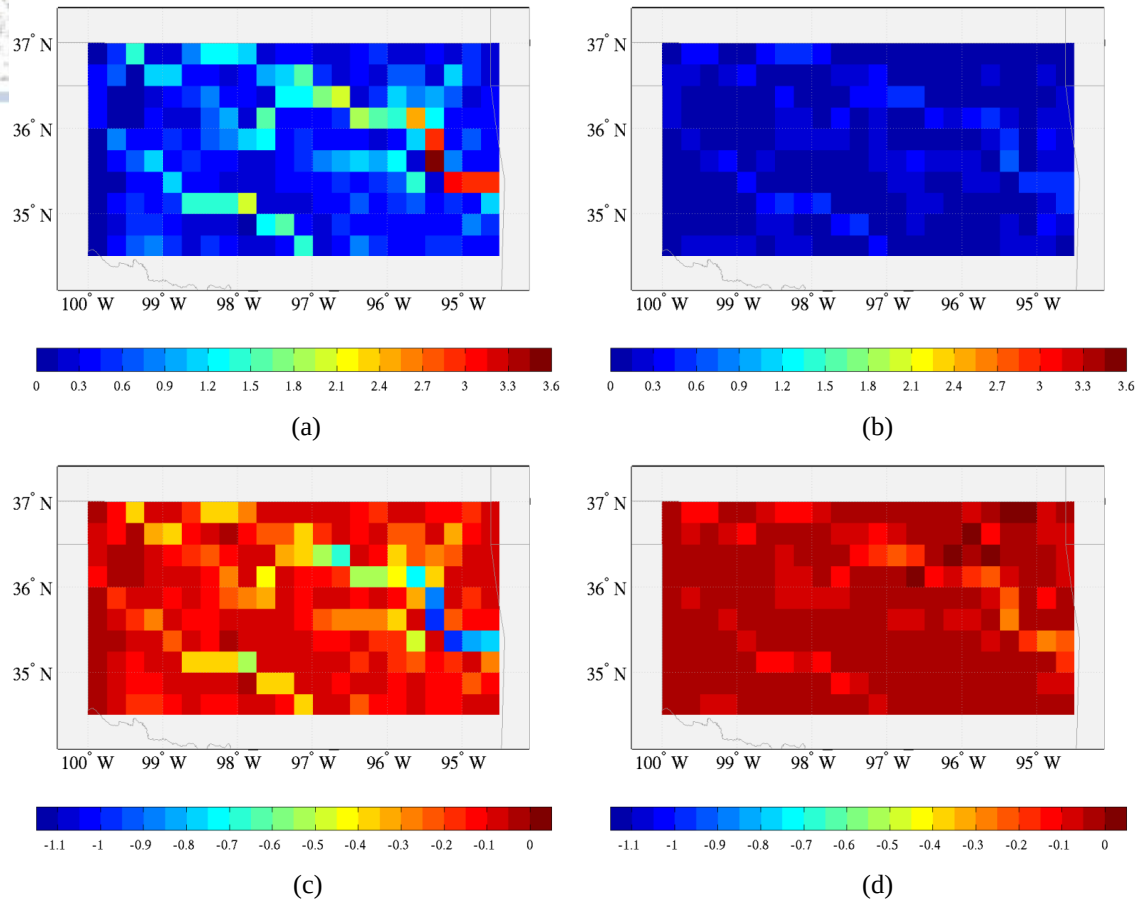


Figure 3.12. Upper panels: Differences in mean root zone soil moisture between CLM-MOD2 and CLM-CTRL runs for (a) 2005 and (b) 2006. Lower panels: The same as upper panels, but for LST at 18 UTC (Terra overpass).

3.6 Chapter summary and discussion

The current study has presented evidence that rivers and surface water affect soil moisture and temperature at the regional scale through the interaction with groundwater. This process is suggested by the analysis of soil moisture data in Oklahoma (Oklahoma Mesonet) and confirmed by a numerical experiment. The numerical experiment was based on CLM (version 3.5), which was modified to include a new parameterization of stream-groundwater interactions based on a groundwater recharge term that relates to the river water flows. This first exploratory study identifies two possible formal dependencies of the groundwater recharge on river water amount that can be implemented in regional climate models. The modified CLM model increases the spatial variability of soil moisture,

affecting the surface fluxes and surface temperature, reducing significantly the bias in soil moisture and surface temperature related to the river distribution. It is effective in producing wetter soil in correspondence to streams and rivers; therefore, it is potentially capable of reproducing effects like, for instance, the lagged correlation between winter snow over mountains and spring soil moisture in the surrounding planes. Our approach differs from the river conductance method described in Miguez-Macho et al. (2007) in that it is formulated only in terms of regional scale variables. Moreover, our analysis is conducted on a 0.25-degree grid resolution, at which regional climate models are commonly run and that high-resolution global climate models are approaching to. Therefore, the newly developed parameterization is directly applicable to these climate/regional model applications.

Results of this study were based on a limited domain associated with moderate orography and river network complexity. More extensive comparison with observations is needed to reduce the uncertainty in the parameterization and the involved closure parameters. Probably, a limiting value of the maximum induced flux is necessary if the model is to be implemented over a greater and/or more complex terrain domain, where the volume of river water could be much larger than the values produced with our simulations over the Oklahoma domain. In fact, the domain over which the set of MESONET observations are collected is not a closed system from a hydrological point of view. The inflow of river discharge entering the study-domain (mainly the Arkansas River) is not accounted in the calculations since it is not generated within the study domain. This is a limiting factor for the accuracy of our study, because the values of river water computed by the RTM scheme are underestimated. This underestimation in the river water accumulation values is expected to have affected the values of the involved parameters. However, the parameterization itself still holds as the demonstrated dependency of river water on the contributing area, besides the bias from the neglected incoming flow, is determined by the rainfall-runoff and evaporation budget inside the domain. We don't account for the possible effect of water management on soil moisture explicitly. However, it is likely that agriculture takes place across most of the domain, and probably we would expect more irrigation taking place in the West and drier areas of the study region, which can only moderate the findings of the study. However, in our framework, the water

management process could be considered a local redistribution in time (low pass filter) of the river water feeding the soil in the neighbouring grid cells. Therefore, in a way, we may be actually accounting for some aspects of this process in our analysis, but it would be difficult to quantify and to separate it from the natural effect due to the river-groundwater interactions. This could be the subject of a future study. Data from the ESA SMOS mission, as discussed in chapter 1, can provide a suitable spatial resolution and coverage to refine our new parameterization at global scale. Another possibility to refine our parameterization will be by adopting a fine scale distributed hydrological model that resolves run-on and infiltration of river water explicitly. In this way we could overcome the problems associated with the limited amount of observations. In our simulation the sensitivity is underestimated because of the surface forcing, therefore coupling with an atmospheric model would allow to quantify also the feedbacks on the atmospheric circulation and to fully explore the implications of the river-groundwater interactions for regional scale modelling. These aspects may be explored in future studies. As a final remark we mention that the inclusion of the reinfiltration of river water into the land surface model improved the comparison with satellite data, therefore our new parameterization may be adopted in data assimilation procedures of satellite data in land surface models to provide more realistic results.

Chapter 4 Data assimilation techniques in NWP

4.1 Introduction

As noted in the Introduction in Chapter 1, Land Surface Models (LSMs) are considered as viable tools for the definition of the land surface conditions that are required in the atmospheric NWP models. The accuracy of the performance of the LSMs depends on the quality of (1) the initial conditions of the soil state (i.e., soil moisture, soil temperature and snow cover), (2) the atmospheric forcing (i.e., precipitation, radiation budget, air temperature, wind, humidity, and pressure), commonly provided by the NWP models, (3) information of the topography, soil properties and land cover, and (4) the internal model parameterization schemes.

Ongoing research investigates the potential benefits of assimilating ground-based and remotely sensed data into advanced LSMs coupled with atmospheric models. Following this trend, one of the major goals of this thesis was to investigate the potential of improving NWP forecasts through off line forcing of an atmospheric model's land surface scheme using remotely sensed precipitation fields. Two major studies were performed in two different areas of the world using different models and data sources, so that our investigation is not limited with respect to location, model and precipitation data-source. As part of the first study, rainfall data from the US WSR-88D radar network (NEXRAD) are used to force the land surface model of the POSEIDON NWP system replacing the model-generated fields. The second study used instead the CMORPH satellite rainfall retrievals (see par. 2.2) to force the Noah land surface model of the Weather Research and Forecasting (WRF) system.

The assimilation technique enables us to investigate the sensitivity of the systems' performance in varying soil state conditions. By providing more accurate precipitation information (derived from radar or satellite observations) to the land surface model, we expect a more realistic distribution of the soil moisture content leading to improved parameterization of the Bowen ratio, i.e., the partition of energy between the sensible and the latent heat fluxes at the surface. These better surface conditions affect the land-atmospheric interactions through convection resulting to a better atmospheric predictability. To assess the effectiveness

of the approach, numerical experiments were carried out with the aid of two major mesoscale atmospheric models, showing a reasonable improvement of the convective precipitation forecasts as confirmed by the radar- or satellite- based rainfall observations. The NWP systems and remote-sensing data used in the current study are briefly described in section 4.2. Section 4.3 describes the numerical experiments performed as well as a detailed analysis of the results for the USA/POSEIDON study. Section 4.4 discusses the EUROPE/WRF study. A comparison and conclusions are provided in section 4.5.

4.2 Numerical models and datasets

4.2.1 The POSEIDON weather forecasting system

The POSEIDON weather forecasting system (Papadopoulos et al. 2002) has been developed in the framework of the project "Monitoring, forecasting and information system for the Greeks seas" and for more than 15 years provides daily 72-h weather forecasts for the Mediterranean basin and surrounding countries (www.poseidon.hcmr.gr). Its central component is the SKIRON/Eta model (Kallos et al. 1997), which is a modified version of the Eta/NCEP model. The convective effects are parameterized using the revised Betts–Miller–Janjic (BMJ) convective scheme (Betts 1986; Betts and Miller 1986; Janjic 1994), while for the grid-scale precipitation a simplified explicit cloud water scheme (Zhao and Carr 1997) is used. For the simulation of the surface processes, Eta uses a two-layer soil model developed at Oregon State University (OSU), including surface hydrology with a vegetation canopy (e.g., Chen et al. 1996). The radiation package used in the model was developed at the Geophysical Fluid Dynamics Laboratory based on the work of Fels and Schwarzkopf (1975) and Schwarzkopf and Fels (1991). More details on the model dynamics and physics packages can be found in previous studies (e.g., Janjic 1984; Mesinger et al. 1988). The POSEIDON/Eta shares the same physics packages as the Eta/NCEP Model with some modifications concerning the use of six layers in the soil model component incorporating a fine dataset of soil textural classes and land cover, and the introduction of the slopes and

azimuths of the sloping surfaces in the calculations of the incoming solar radiation on inclined surfaces (Papadopoulos et al. 1997).

4.2.2 The Weather Research and Forecasting (WRF) system

The Weather Research and Forecasting (WRF) model is a next-generation mesoscale NWP system designed for both atmospheric research and operational forecasting applications (for more details, please visit https://www.mmm.ucar.edu/weather-research-and-forecasting-model). It features two dynamical cores, namely the Non-hydrostatic Mesoscale Model (NMM) core and the Advanced Research-Weather (ARW) core, a data assimilation system, and a software architecture supporting parallel computation and system extensibility. The model serves a wide range of meteorological applications across scales from tens of meters to thousands of kilometres. The effort to develop WRF began in the latter part of the 1990's and was a collaborative partnership of the National Centre for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (NOAA; represented by the National Centres for Environmental Prediction (NCEP) and the Forecast Systems Laboratory (FSL)), the (then) Air Force Weather Agency (AFWA), the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA).

The WRF-NMM, version 3 (Janjic, 2003), was chosen for the specific study as its structure is also based on the NCEP/Eta model. WRF-NMM uses the vertical sigma-pressure coordinates following topography, while in the horizontal layer the grid-geometry is based on the Arakawa E-grid. WRF-NMM includes multiple options with respect to the parameterizations of the physical processes. For the purposes of the thesis and in order to be as close as possible to the POSEIDON system, the following parameterizations were used: the Betts-Miller-Janjic (BMJ) convection parameterization scheme (Betts and Miller, 1986; Janjic, 1994; 2000), the Mellor-Yamada-Janjic (MYJ) scheme for the reproduction of the boundary layer processes (Janjic 1996a; 2002) while there is a specific separation of the boundary layer that is solved based on the Janjic scheme and the similarity theory (Janjic, 1996b; Chen et al., 1997), the Ferrier scheme for the simulation of microphysical processes

(Ferrier et. al., 2002), as well as the USA Geophysical Fluid Dynamics Laboratory (GFDL) schemes for the calculation of shortwave- (Lacis and Hansen, 1974) and longwave radiation (Fels and Schwarzkopf, 1975; Schwarzkopf and Fels, 1991). Among the various options for the land surface modelling scheme, the Noah LSM was used, which outputs soil moisture and soil temperature in 4 different layers (0-7 cm, 7-28 cm, 28-100 cm, 100-255 cm). It is worth noting here that WRF-NMM, specifically with the parameterizations mentioned above, has been the major operational NCEP model for weather forecasting in USA for the past decade.

4.2.3 Forcing and assimilation data

For the initial meteorological conditions (geopotential height, wind components, temperature and specific humidity), the ECMWF reanalysis gridded data were used for both models/studies on a 0.5-degree horizontal grid increment. Regarding the vertical grid, 11 standard pressure levels (1000, 925, 850, 700, 500, 400, 300, 250, 200, 150 and 100 hPa) were used for the POSEIDON runs, while 17 standard pressure levels (surface, 1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20 και 10 hPa) were used for the WRF runs. These data were then interpolated at the model grid points using optimal interpolation analysis. The boundary conditions were linearly interpolated at each model time step from the ECMWF data available every 6 hours. For each model grid point we calculate topography, soil and vegetation types, slopes and azimuths of the sloping surfaces, utilizing high resolution datasets. Namely, the topographic and vegetation datasets, both available from USGS at 30x30 arc sec resolution, and the UNEP/FAO dataset after being converted from soil type to soil textural ZOBLER classes at 2x2-min resolution (Papadopoulos et al. 1997).

The rainfall data used for assimilation in the land surface scheme of the NWP systems are high-resolution remotely-sensed data, coming either from weather radars (NEXRAD data in the case of USA/POSEIDON study) or from satellite sensors (CMORPH data in the case of the European/WRF study). Please see paragraph 2.2 for a detailed description of the NEXRAD and CMORPH products. The choice of these datasets was based primarily on availability and high-resolution coverage of the areas under study.

4.3 USA/POSEIDON study

4.3.1 Study area

Rainfall data from the US WSR-88D (NEXRAD) radar network and lightning data from the US National Lightning Detection Network (NLDN) were used in the current study to identify areas of convective activity over continental USA for the 2004 summer period. Time-series of 6-hourly lightning-to-rain ratio, defined as total number of flashes to total accumulated precipitation occurring within a box defined by 80° W, 115° W, 25° N and 45° N, were plotted (graphs not shown here) and helped to identify the temporal domains of intense convective activity within a 4-month period of May to August 2004. In addition, monthly Hovmöller diagrams were constructed of NEXRAD rainfall rates overlapped by the number of lightning flashes for the aforementioned domain. Spatial domains of intense convective activity were then recognized for the periods identified in the time-series of lightning-to-rain ratio, as longitudinal zones with increased precipitation rates overlapping with pronounced lightning activity. Finally, 6-hourly geographical maps of rainfall and lightning activity were constructed for the prescribed temporal and spatial domains. After careful examination of these maps, a 48-hour mesoscale convective system occurring over parts of Kansas, Missouri, Oklahoma and Arkansas from 23 July 2004 at 12 UTC to 25 July 2004 at 12 UTC was considered suitable for performing our numerical experiments. The full domain used in all numerical experiments described in the following section is shown in Fig. 3.1, while the box denotes the storm area under investigation (confined by 90° W, 100° W, 34° N and 41° N).

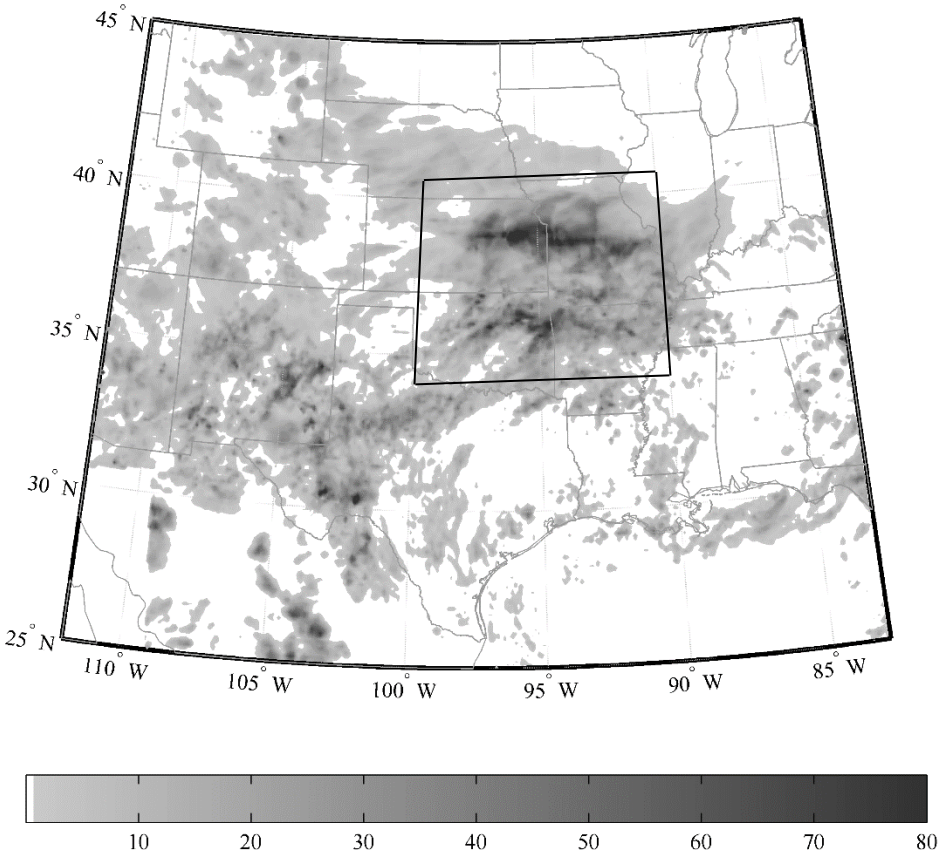


Figure 4.1. Map of total rainfall accumulation (in mm - shaded contours) for the 48-h period ending at 12 UTC on 25 July 2004 for the full domain used in the numerical experiments. The box denotes the area under investigation.

4.3.2 Experimental setup

As discussed above, soil moisture, soil temperature and precipitation are considered very important parameters in land-atmosphere interactions, having significant impact on surface water and energy budgets and significantly affecting the atmospheric variability and weather prediction. With this in mind, numerical experiments were performed to assess the sensitivity and impact on performance of the POSEIDON weather forecasting system (especially the convective precipitation forecasting skill) to the utilization of the WSR-88D rainfall estimates as land surface precipitation forcing instead of the model-generated fields.

For the specific numerical experiments, the POSEIDON system was implemented over the data-rich region of the continental US covering an 84-hour simulation period (from 22 July 2004 at 00 UTC to 25 July 2004 at 12 UTC) with a horizontal grid increment of 0.1x0.1 deg and a time step of 36 seconds. Three experiments were designed to assess the effectiveness of the land surface data forcing scheme. In the control run (named CTRL hereafter) the initial soil conditions were computed at the 6 soil layers (currently defined at the depths of 5, 15, 28, 50, 100 and 255 cm) from the ECMWF gridded data through a typical interpolation allowing the OSU land model to carry on as it normally does. In the second experiment (named ASE1), observed rainfall fields derived from a mosaic of hourly raingauge-calibrated radar (WSR-88D) rainfall estimates (Fulton et al. 1998) were used to force the OSU land model replacing the model-generated precipitation fields, while the initial soil conditions were defined in the same way as in the CTRL experiment. In this experiment, the radar data ingestion was performed for the entire simulation period (84 hours) at each model time step after linearly interpolating the hourly rainfall data. While ASE1 allows for investigating the sensitivity of the model-produced rainfall fields to improved representation of the precipitation forcing of the land surface model, it cannot be used as a direct forecasting scheme. Thus, a third experiment was designed (named ASE2) to investigate the effectiveness of radar rainfall data forcing in 48 hours of consequential simulations. The ASE2 experiment is schematically described in Fig. 4.2; it consists of three separate model runs accounting for different time periods of rainfall data ingestion prior to the 48-hour forecast/verification period. The ingestion is performed similarly to the ASE1 experiment, but for periods of 12 hours (ASE2-12), 24 hours (ASE2-24) and 36 hours (ASE2-36) all ending at the beginning time of the convective system under consideration (12 UTC on 23 July). Therefore, we can assess the relative contribution of these potentially enhanced surface conditions initializations on the improvement of the mesoscale simulation of the 48-hour forecast/verification period. Since all simulations have a common initialization time (00 UTC on 22 July), the CTRL mode is used in ASE2-12 and ASE2-24 for the period prior to the data ingestion.

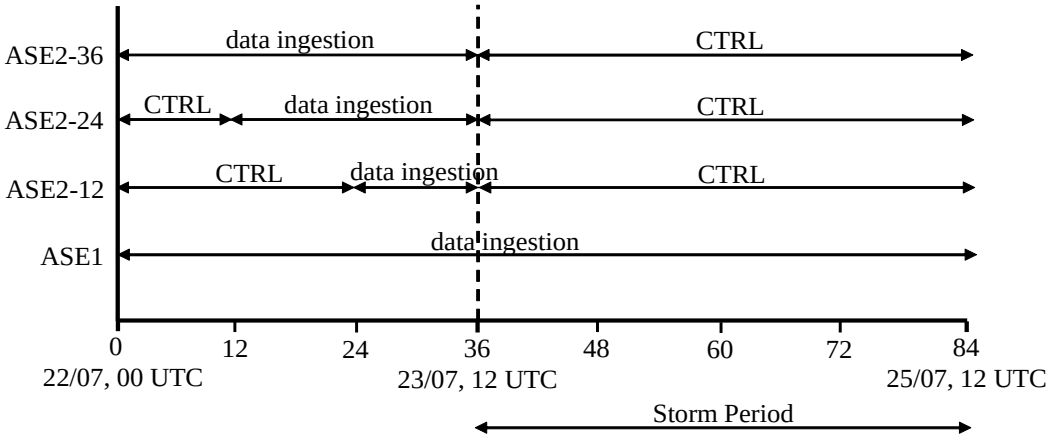


Figure 4.2. Design of ASE1 and ASE2 experiments.

4.3.3 Application of the technique

The accumulated rainfall fields obtained from the three model experiments for a 6-hour period (ending at 18 UTC on 24 July 2004) are shown in Fig. 4.3 (panels b through f denoting CTRL, ASE1, ASE2-12, ASE2-24 and ASE2-36, respectively). The first map (Fig. 4.3a) corresponds to the respective rainfall fields as observed by the NEXRAD radar network. As indicated in Fig. 4.3, after 66 hours of model integration (30 hours within the storm period) the CTRL run (panel b) produces some of the patches of the 6-hourly accumulated precipitation fields observed at the radar rainfall observations (panel a). However, the implementation of ASE1 (panel c) and ASE2 (panel d through f) experiments seems to bring the modelled rainfall fields closer to the ground truth (observed data rainfall fields), by enhancing rain rates in the regions where they were underestimated in the CTRL run (e.g., over Missouri) and limiting them in the regions where they were overestimated (e.g., over Oklahoma). Especially for the 6-hour period under investigation, the model's rainfall representation exhibits improvements not only for the simulation implemented with continuous data ingestion (ASE1) but also for the shorter cases of 24- and 36-hours of data ingestion prior to the 48 hours of free model run (ASE2-24 and ASE2-36, respectively).

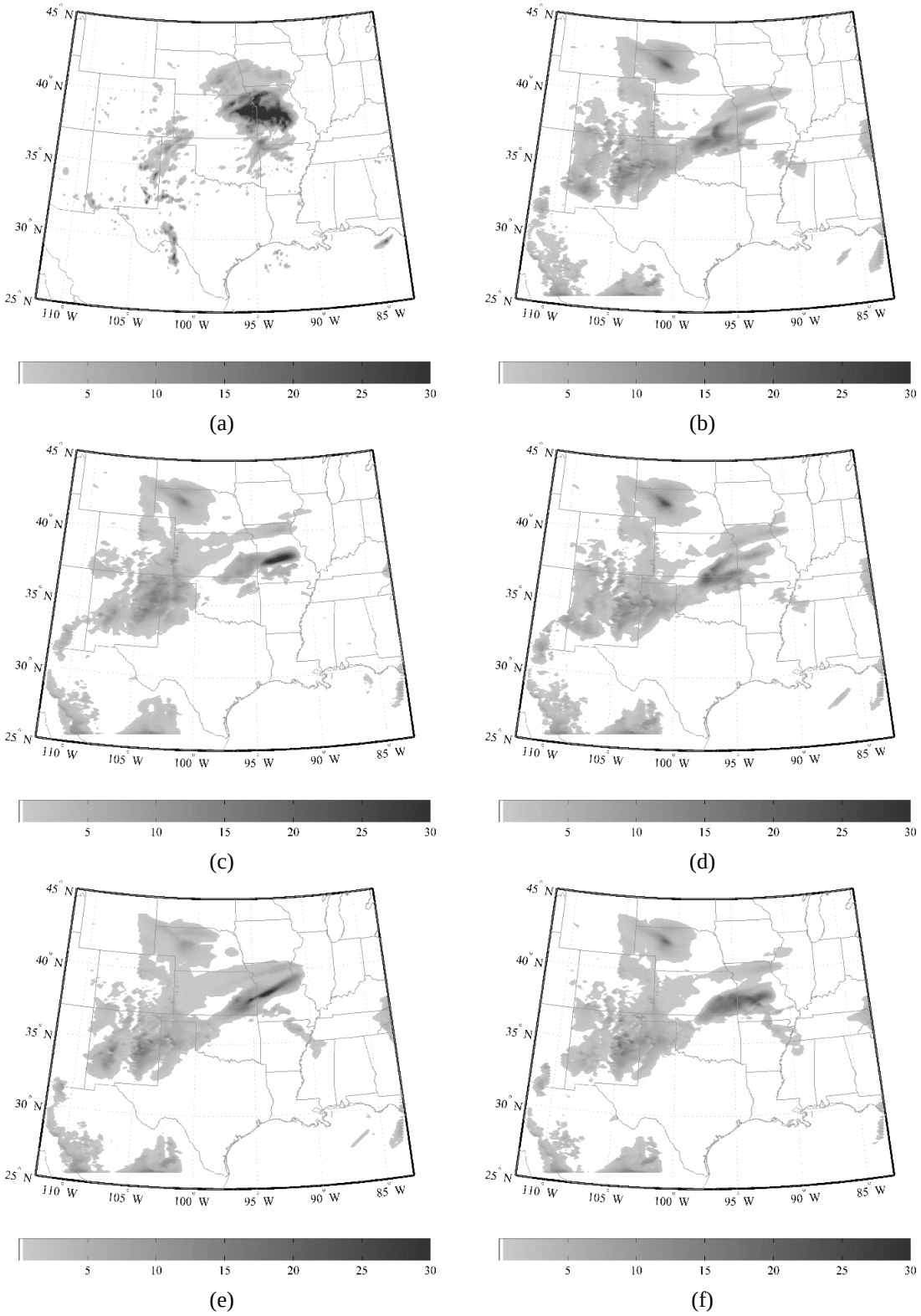


Figure 4.3. Rainfall fields (in mm) as observed by the NEXRAD radar network (a), and as predicted by CTRL (b), ASE1 (c), ASE2-12 (d), ASE2-24 (e) and ASE2-36 (f) experiments, for a 6-h period ending at 18 UTC 24 July 2004.

4.3.4 Evaluation methodology

To provide quantitative estimates of the improvement of the model performance in convective precipitation forecasting, the model precipitation forecasts are verified against the measured NEXRAD rainfall fields for the storm area under investigation (90° W - 100° W, 34° N - 41° N). For this reason, precipitation from the radar pixels was interpolated to each model grid point using bilinear interpolation:

$$R = \frac{\sum_{k=1}^{4} w_k \cdot R_k}{\sum_{k=1}^{4} w_k}, \qquad \textit{(Equation 4.1)}$$

where R_k are the values at the four radar pixels surrounding each model grid point, while the weighting factor w_k is taken as the reverse squared distance, thus making nearer points more influential.

The verification scores used in this work are derived using the contingency table approach (Wilks 1995). This is a two-dimensional matrix where each element counts the number of occurrences in which the observations and the model forecasts exceeded or failed to reach a certain threshold for a given forecast period. The table elements are defined as: A-model forecast and observation exceeded the threshold; B-model forecast exceeded the threshold but observation did not; C-model forecast did not reach the threshold but observation exceeded it; and D-model forecast and observation did not reach the threshold. Considering the above elements, the forecast skill can be measured by evaluating the bias score (BS) and the threat score (TS). The bias score is defined by:

$$BS = \frac{A + B}{A + C}, \qquad \textit{(Equation 4.2)}$$

where BS defines the ratio of the number of occurrences that model forecasts exceed a specified threshold versus the respective number for observations. The TS score is defined as:

$$TS = \frac{A}{A + B + C}, \qquad \textit{(Equation 4.3)}$$

Computing the bias and threat scores, a measurement of the model accuracy on the frequency of occurrences at or above a certain precipitation threshold amount can be revealed. Consequentially, at given thresholds the bias score can represent a systematic overestimation (when BS>1) or underestimation (when BS<1), and the threat score can represent poor forecasts (when TS≈0) or the perfect forecasts (when TS=1).

Another widely used score for verifying precipitation forecasts is the Heidke Skill Score, HSS (Heidke 1926). It is computed based on the contingency table elements from the expression:

$$HSS = \frac{2 \cdot (A \cdot D - B \cdot C)}{(A + C) \cdot (C + D) + (A + B) \cdot (B + D)} \qquad \textit{(Equation 4.4)}$$

HSS is another measure of correspondence between the estimate and the reference (see for example Barnston, 1992). It is defined such that a perfect forecast would be scored as 1, a random forecast would have an expected score of 0, and forecasts having fewer hits than would be expected by chance would have negative scores. HSS combines the effects of probability of detection, false alarm rate and occurrences by chance. Thus, it is a better measure as compared to using just probability of detection and false alarm rate.

Since the aforementioned statistical measures do not use the magnitude of the precipitation errors, they are not strictly influenced by the variability of forecasting error. To measure the magnitude of the difference between model forecast and observed precipitation we calculate the root mean square error (RMSE) as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{NOBS} (MP_i - OP_i)^2}{NOBS}}, \qquad \textit{(Equation 4.5)}$$

where MP_i and OP_i are the model-estimated and the observed precipitation, respectively, and NOBS is the total number of observations at a specific location reaching or exceeding a certain threshold amount. Combining these statistical criteria, we attempt to provide a comprehensive evaluation of the model performance. For example, a greater TS will

represent a significant model improvement only if it is accompanied by BS and HSS with values close to one and a lowering RMSE.

4.3.5 Evaluation results

The results of the statistical analysis are presented in Figs. 4.4 and 4.5. Fig. 4.4 refers to the 6-hour period ending at 18 UTC on 24 July, while Fig. 4.5 includes the statistical evaluation for the entire storm period (48-hour period ending at 12 UTC on 25 July 2004). The left panels in each figure depict the calculated values of the statistical scores for the three numerical experiments, while in the right panels the relative differences (in %) of ASE1 and ASE2 experiments versus CTRL score values are shown. A noteworthy improvement is noted in all statistical scores when radar rainfall data are continuously ingested into the model (ASE1 experiment). For the specific 6-hour period, there is a mean relative improvement of about 20% in the bias score, which is accompanied by 100% improvement in the threat and Heidke skill scores for almost the entire span of precipitation thresholds, and a mean relative RMSE reduction of about 15%. Less but still distinctive improvement is exhibited in the respective statistical scores of the ASE1 experiment for the entire storm period. We cannot infer whether low or high precipitation thresholds, and thus weak or intense rainfall rates, are favoured by the continuous ingestion of radar data into the land surface model; for TS and HSS, the relative improvement increases with increasing rain accumulation threshold, while the opposite pattern is observed for BS and RMSE.

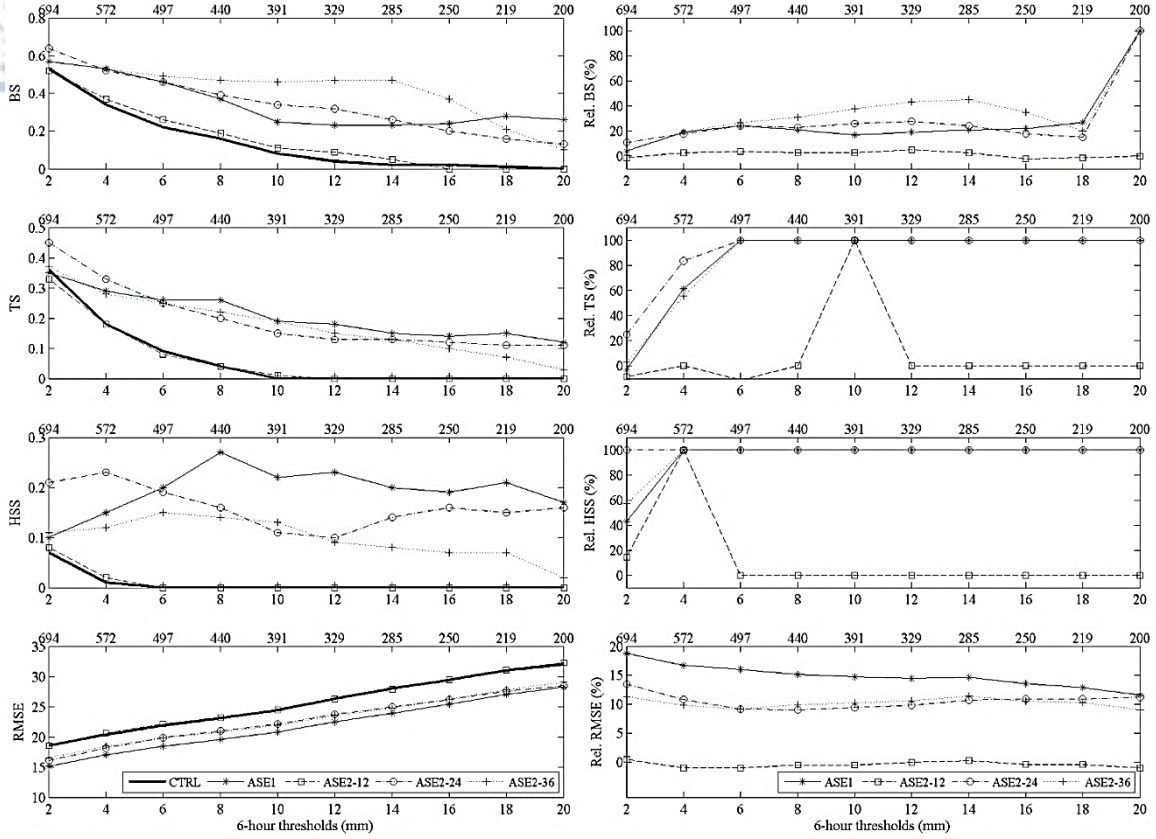


Figure 4.4. Statistical evaluation of the CTRL, ASE1 and ASE2 experiments against radar rainfall fields for the 6-hour period ending at 18 UTC 24 July 2004. Left panels: Calculated values of BS, TS, HSS and RMSE for the three experiments. Right panels: Relative differences of ASE1 and ASE2 experiments versus CTRL for all statistical scores. Numbers above each panel denote the total number of observations reaching the corresponding threshold value.

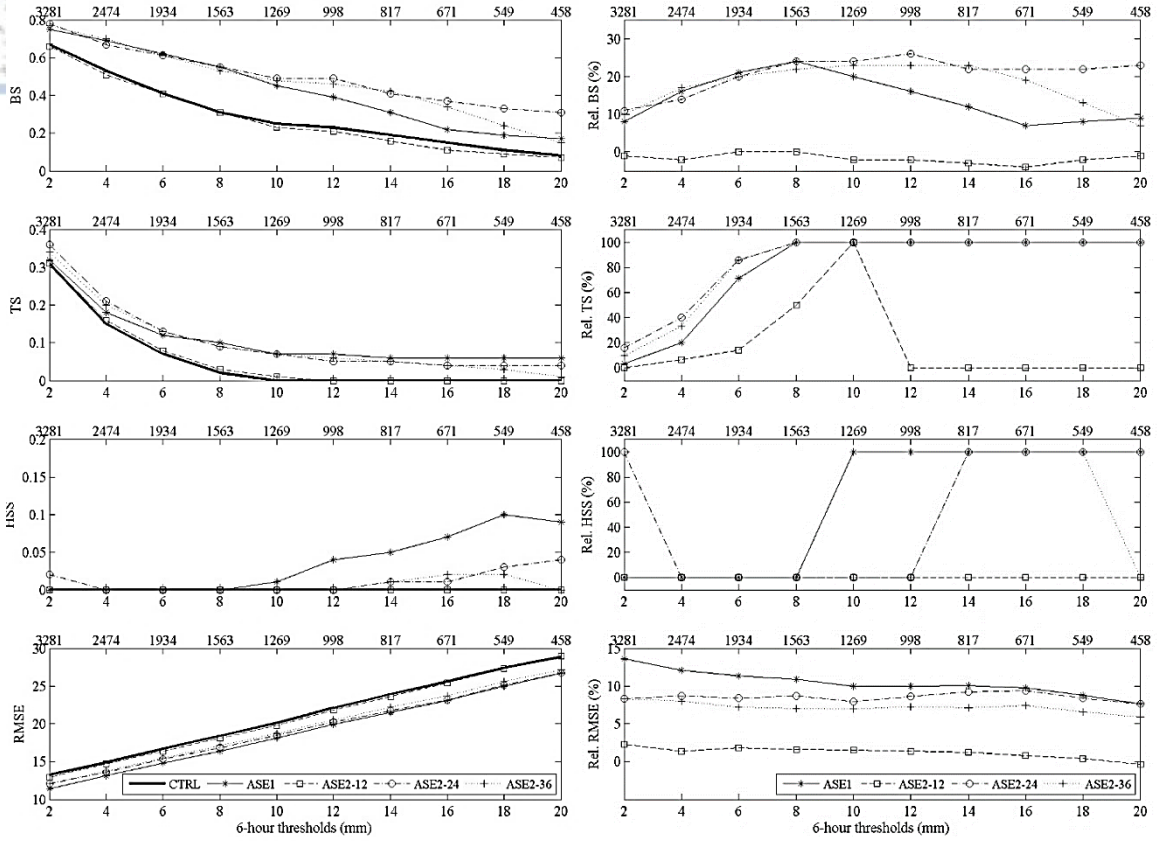


Figure 4.5. As in **Figure 4.4**, but for the entire storm (forecast) period.

The results of the ASE1 experiment are quite promising with respect to the improved response of the modelled precipitation fields to the ingestion of remotely sensed rainfall data into the land surface scheme of the NWP model. However, the continuous assimilation mode cannot be used as an operational tool for getting real-time precipitation forecasts. As mentioned before, the ASE-2 experiment investigates this possibility through the implementation of short-time integration periods prior to free-run simulations that could be used for forecasting purposes. The results shown in Figs. 4.4 and 4.5 show improvement for both 6- and 48-hour forecasts in the cases of 24 and 36 hours of preceding radar data integrations when compared to the corresponding CTRL forecasts. The ASE2-12 experiment (12 hours of integration prior to the free run) does not seem to provide any improvement to the forecasted rainfall fields and the statistical scores associated with this experiment seem to closely follow the respective scores of the CTRL run.

Further evaluation on the ASE2-24 and ASE2-36 experiments reveals that there are cases where these experiments perform slightly better than the continuous data ingestion mode (ASE1 experiment). Indeed, for both the 6-hour period ending at 18 UTC on 24 July and the entire 48-hour storm period, the relative improvement versus the CTRL run of the bias score for the ASE2-24 and ASE2-36 experiments is higher than the respective improvement of the ASE1 run, especially for higher precipitation thresholds. For the TS score, the three experiments exhibit about the same relative improvement versus the CTRL simulation, whereas the HSS and RMSE evaluation shows that the continuous data ingestion experiment performs better than the reduced-time runs.

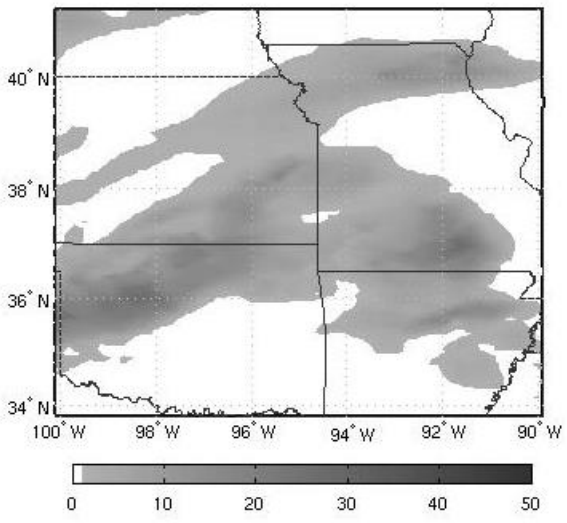
Moreover, a comparison between ASE2-24 and ASE2-36 experiments is substantiated by the need to define the time duration most suitable for applying our off-line land model data forcing technique, based on the criterion of best forecasting performance for the shortest integration period possible. After examining Figs. 4.4 and 4.5, we note that the 24 and 36 hours of radar data ingestion prior to the free model simulation do not exhibit significant differences with respect to the relative improvement that both impose on the CTRL run forecasts. Although for the 6-hour period under investigation the ASE2-36 run seems to have an overall better performance than ASE2-24, this is not the case for the entire storm period; ASE2-24 and ASE2-36 maintained similar scores in the cases of TS and HSS. Thus, a 24-hour period of assimilation can be considered adequate to improve the model's precipitation forecasting for the convective system examined in this study.

4.3.6 Physical interpretation

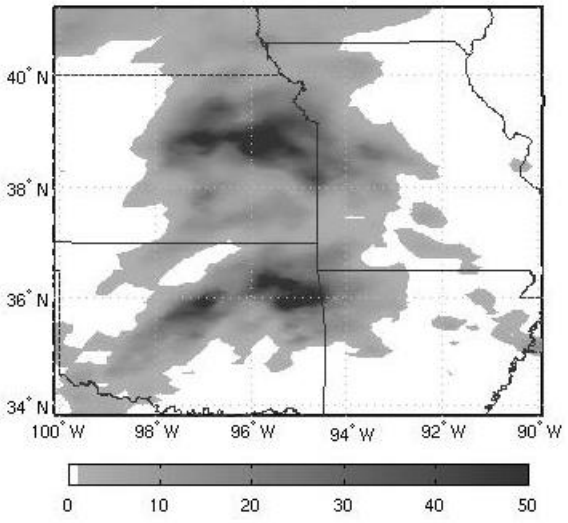
The technique used in the current study is substantiated by the need for advancing our understanding on the processes and feedbacks that modulate the local-scale spatiotemporal interactions between soil state and precipitation. The use of radar rainfall data in place of the modelled precipitation fields as input to the land surface scheme may result to a more accurate estimation of the soil moisture content. Actually, introducing an improved spatiotemporal distribution of soil moisture and through the partition between the sensible and latent heat fluxes, a direct impact on the spatiotemporal distribution of surface heating

should be established. This heterogeneous surface heating is a critical factor in modifying the mesoscale atmospheric circulations associated with convection. Entekhabi (1995) showed that wet soil conditions force larger equivalent potential temperature, greater cloudiness and precipitation potential. In this study we attempt to demonstrate that the link between surface state and the atmospheric hydrologic cycle involves the atmospheric boundary layer and consequentially precipitation even at the short-term periods resolved by mesoscale models.

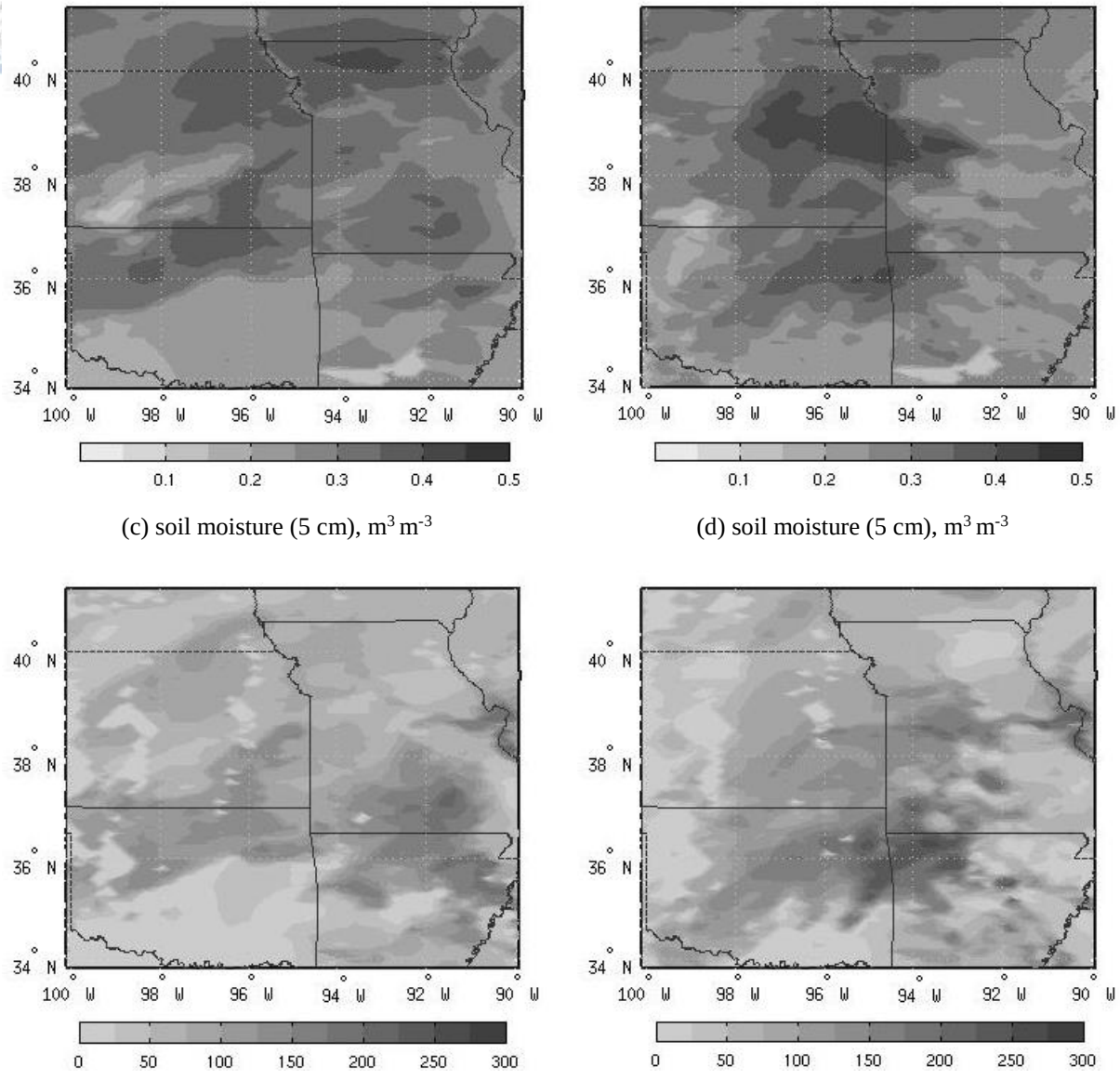
In this section we assess the strength of the above feedback mechanisms on the basis of the numerical experiments described in the previous sections. Figs. 4.6 and 4.7 facilitate a graphical investigation of the interactions between the aforementioned surface properties, i.e. rainfall, near surface soil moisture and temperature (at 5 cm), latent and sensible heat fluxes, on the basis of their differences between CTRL and ASE1 experiments at 14 UTC on 24 July 2004 (26 hours after the storm initiation) for the area confined by 90° W, 100° W, 34° N and 41° N.



(a) CTRL model predicted rainfall, mm



(b) NEXRAD rainfall, mm



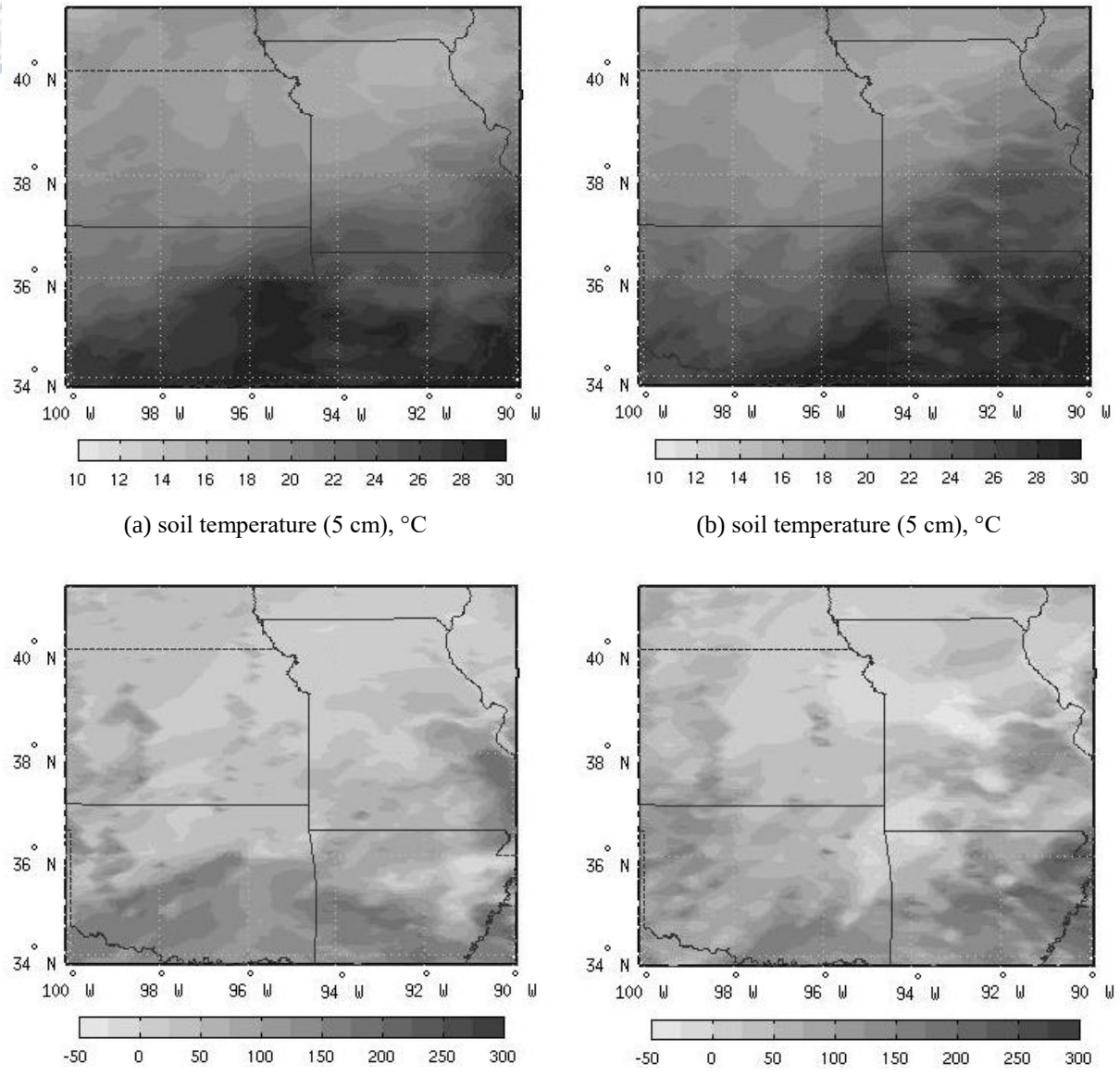
(c) soil moisture (5 cm), $m^3\ m^{-3}$

(d) soil moisture (5 cm), $m^3\ m^{-3}$

(e) latent heat flux, $W\ m^{-2}$

(f) latent heat flux, $W\ m^{-2}$

Figure 4.6. Surface fields of various land-air interaction properties at 14 UTC 24 July 2004 for the area under investigation. Left and right panels show CTRL and ASE1 experiments, respectively. The top panels show 12-hour accumulated rainfall (mm) for the period 02-14 UTC on 24 July 2004 as predicted by the model (panel a) and observed by NEXRAD (panel b). The model-predicted land surface parameters shown in this figure are soil moisture at 5 cm (panels c and d) and latent heat flux (panels e and f).



(a) soil temperature (5 cm), °C

(b) soil temperature (5 cm), °C

(c) sensible heat flux, W m^{-2}

(d) sensible heat flux, W m^{-2}

Figure 4.7. As in **Figure 4.6** but showing soil temperature at 5 cm (panels a and b) and sensible heat flux (panels c and d).

The 12-hour (2-14 UTC) accumulation of rainfall fields ingested in the OSU land model (Fig. 4.6b) appear to have a strong impact on the predicted near surface soil moisture content (Fig. 4.6, panel d as opposed to panel c). Soil moisture modulations seem to follow the spatial rainfall patterns ingested in the land model during the ASE1 experiment. The feedback is strongest in the areas of eastern Kansas and western Missouri, where intense

rainfall causes significant increase in near surface soil moisture (at the order of 0.05 to 0.1). The modified soil moisture fields appear to further influence the latent heat flux spatial distribution (Fig. 4.6, panel f as opposed to panel e). Latent heat flux is higher (by about 50 to 75 Wm^{-2}) in the areas where soil moisture shows the highest increase from CTRL to ASE1 experiment, e.g. western Missouri. The soil temperature (Fig. 4.7, panel b as opposed to panel a) and sensible heat flux (Fig. 4.7, panel d as opposed to panel c) spatial distributions are also affected by the modulation of soil moisture fields, but the sign of these changes is weakly defined from the respective panels. Nevertheless, we note slightly colder temperatures and reduced sensible heat fluxes in areas where precipitation is ingested due to radar data.

To quantify the impact of the assimilation technique on the prescribed land-air interaction parameters, the distributions of the relative differences between the ASE1 and CTRL experiments are plotted in Fig. 4.8, separately for the grid points where the 12-hour NEXRAD accumulated rainfall is greater than the corresponding 12-hour rainfall predicted by the model CTRL experiment, and vice versa. Relative difference (RD) is defined as:

$$RD_V = \frac{V_{ASE1} - V_{CTRL}}{V_{CTRL}}, \qquad \textit{(Equation 4.6)}$$

where V represents the different land surface parameters (near surface soil moisture and temperature, latent heat and sensible heat flux). The mean values of these relative differences (MRDs) are summarized in Table 4.1 that also provides ratios of MRD values of soil moisture, soil temperature, and fluxes of latent and sensible heat to the MRD of precipitation.

Fig. 4.8 and Table 4.1 verify that the areas, where the precipitation ingested in the land model is greater than the respective rainfall passed onto the land model from the atmospheric model, are characterized by increased values of soil moisture and latent heat flux, with the normalized mean relative difference being about 13% for soil moisture and 40% for latent heat flux. Similar values (-15% and -31%, respectively) are obtained in the case of precipitation ingested in the land model being lower than the respective predicted rainfall from the CTRL experiment. The effect on latent heat fluxes exhibits wider spread and more skewed distributions towards larger values (both positive and negative), caused

by the non-linear land surface processes. In terms of soil temperature, the effect is less significant (mean relative difference is below 5%), but in sensible heat fluxes the relative differences are skewed to larger negative values (mean relative differences greater than 15%). Thus, ingested precipitation has the inverse impact on sensible heat flux than latent heat flux. The combined effect of heat fluxes is a noteworthy decrease (increase) in the Bowen ratio in areas where the NEXRAD rainfall ingested to the model is higher (lower) than the corresponding rainfall predicted by the model CTRL experiment, as indicated by the respective frequency distribution shown in Fig. 4.9 (panel a). The mean relative differences are about -65% and 30% for the two cases, respectively. The net effect on the subsequent model prediction of precipitation (2 hours of accumulation: 14-16 UTC) is consequentially high, as shown in Fig. 4.9 (panel b), with mean increase (decrease) in precipitation in the range of 30% in areas where NEXRAD rainfall rates are higher (lower) than CTRL.

Table 4.1. Mean relative differences (MRDs) between ASE1 and CTRL experiments and mean relative differences normalized with the mean relative difference of 12-h accumulated rainfall (MRD/MRD_{rain}) for the parameters shown in **Figure 4.8** and **Figure 4.9**.

	NEXRAD rainfall > CTRL rainfall		NEXRAD rainfall < CTRL rainfall	
	MRD	MRD/MRD_{rain}	MRD	MRD/MRD_{rain}
12-h acc. rainfall (02-14 UTC)	0.82	-	-0.851	-
Soil moisture	0.106	0.129	-0.124	0.145
Latent heat flux	0.334	0.408	-0.267	0.314
Soil temperature	-0.024	-0.029	0.052	-0.061
Sensible heat flux	-0.344	-0.42	0.151	-0.178
Bowen ratio	-0.547	-0.668	0.273	-0.321
2-h acc. rainfall (14-16 UTC)	0.316	0.386	-0.275	0.323

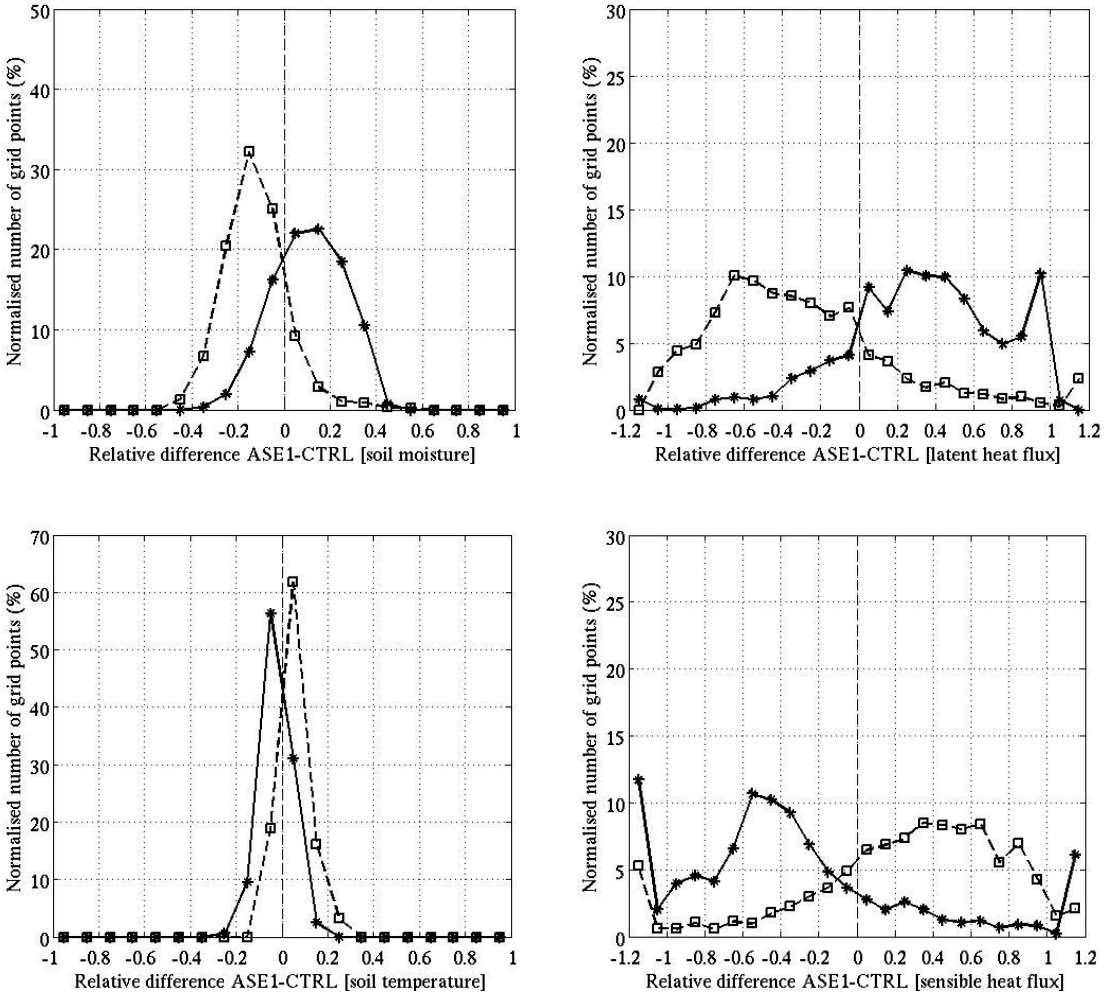


Figure 4.8. Frequency distributions of the relative differences between ASE1 and CTRL experiments for the model-predicted parameters shown in **Figure 4.6** and **Figure 4.7**: near surface soil moisture ($m^3\,m^{-3}$), latent heat flux ($W\,m^{-2}$), near surface soil temperature (°C) and sensible heat flux ($W\,m^{-2}$). For each parameter, the relative difference is calculated for the grid points where the 12-h accumulated rainfall as observed by NEXRAD is higher than the corresponding rainfall predicted by the model CTRL experiment (stars, solid line), and elsewhere (squares, dashed line).

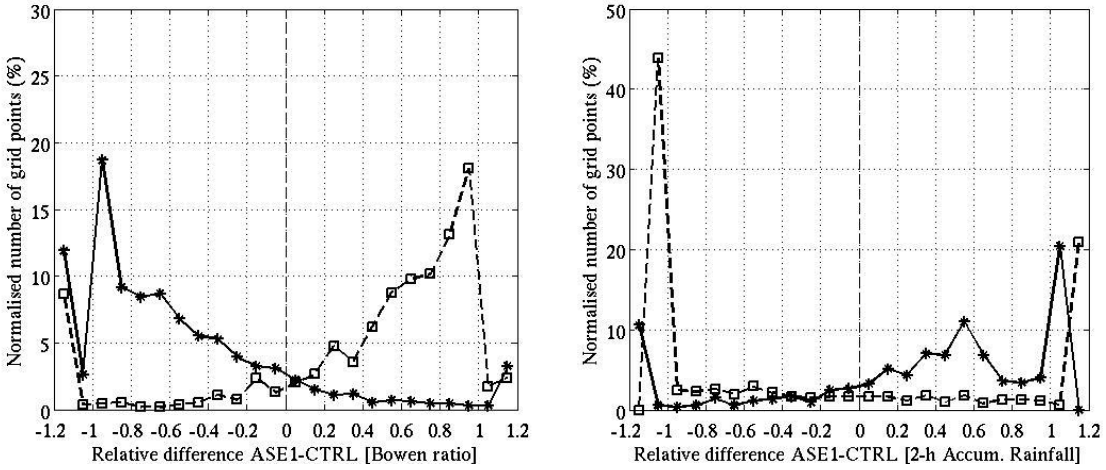


Figure 4.9. As in **Figure 4.8**, but for the Bowen ratio (panel a) and the subsequent accumulated rainfall for the period 14-16 UTC on 24 July 2004 (panel b).

4.4 EUROPE/WRF study

4.4.1 Study area and experimental setup

The same experimental design (as described in 4.3 for the USA/POSEIDON study) was followed for the EUROPE/WRF study. A 96-hour mesoscale convective system of special interest was chosen for the specific study. The specific MCS occurred over parts of Central and South-eastern Europe from approximately 21 to 24 August 2005 and caused a severe flash flood in Romania on the 23rd August 2005 with many casualties and infrastructure damages. The total rainfall accumulation from the CMORPH retrievals for the first 3 days of the cyclonic occurrence (21-23 August 2005) is shown in Fig. 4.10. The depicted area coincides with the area of model integration, while the box denotes the storm area under investigation (confined by 18° E, 31° E, 42° N and 50° N). The development and evolution of the cyclonic system is graphically presented in Fig. 4.11, which shows hourly rainfall instances of the system as captured by CMORPH satellite data (Papadopoulos et al. 2009).

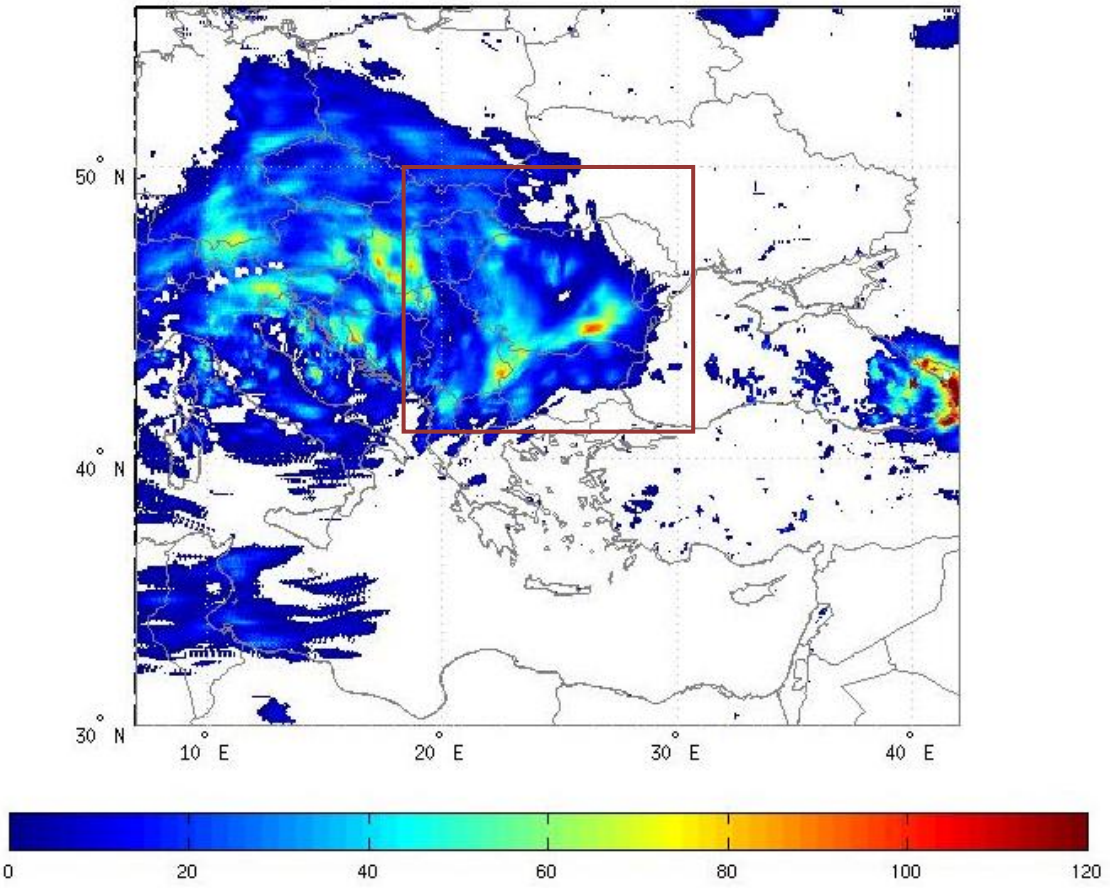
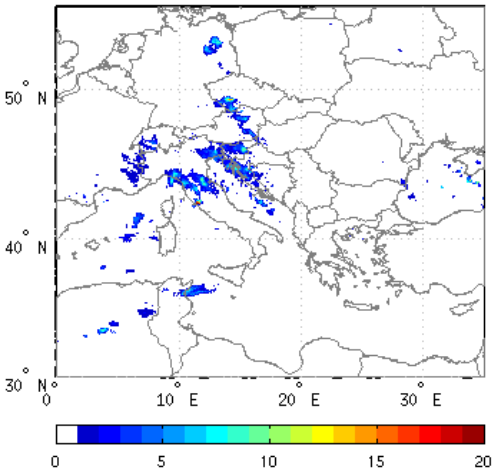
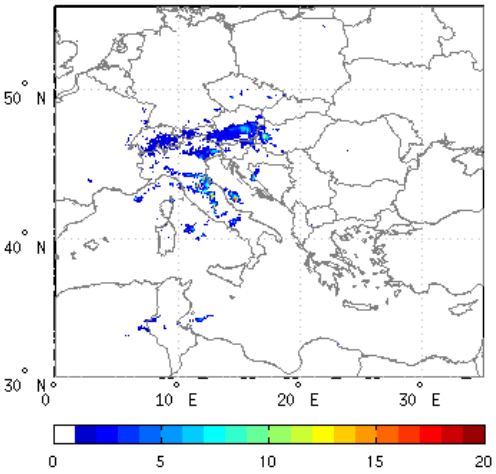


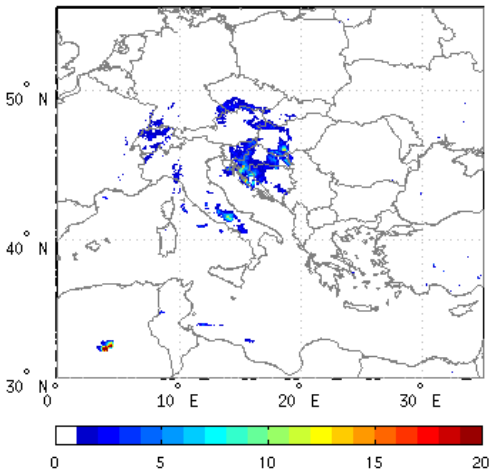
Figure 4.10. Map of total rainfall accumulation (in mm - shaded contours) from CMORPH data for the 72-h period ending at 00 UTC on 24 August 2005 for the full domain used in the numerical experiments. The box denotes the area under investigation.



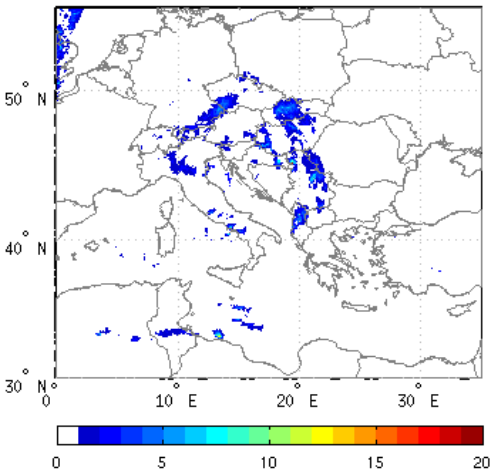
(a) 21 August 2005, 00 UTC



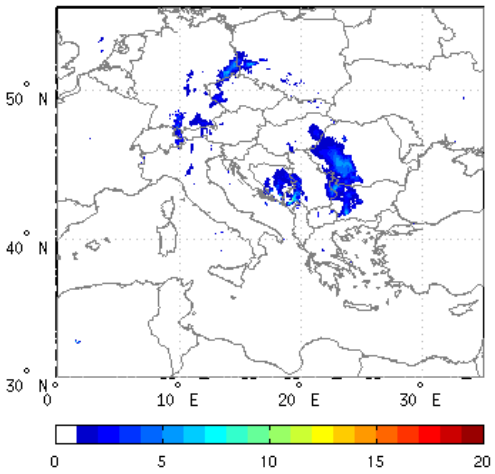
(b) 21 August 2005, 12 UTC



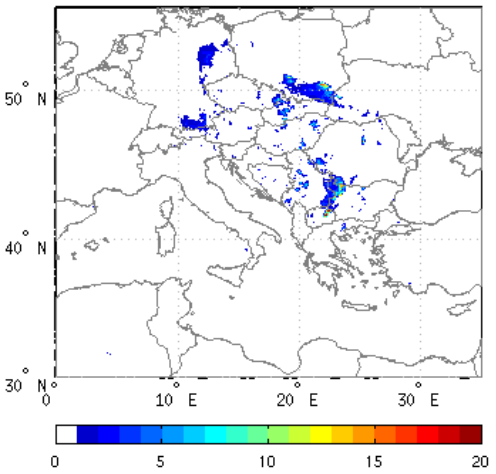
(c) 22 August 2005, 00 UTC



(d) 22 August 2005, 12 UTC



(e) 23 August 2005, 00 UTC



(f) 23 August 2005, 12 UTC

Figure 4.11. Snapshots of hourly accumulated rainfall (mm) from CMORPH satellite data for the hourly periods ending at the times indicated beneath each panel.

As in the USA/POSEIDON study, numerical experiments were performed to assess the sensitivity of the WRF system to the utilization of the CMORPH rainfall retrievals as land surface precipitation forcing instead of the model-generated fields. For the specific numerical experiments, the WRF system was implemented over the region shown in Fig. 10, covering a 96-hour simulation period (from 21 August 2005 at 00 UTC to 25 August 2005 at 00 UTC) with a horizontal grid increment of 9x9 km and a time step of 20 seconds. Two experiments were designed to assess the effectiveness of the land surface data forcing scheme. In the control run (named CTRL hereafter) the initial soil conditions were computed at the 4 LSM soil layers from the ECMWF gridded data through a typical interpolation allowing the Noah land model to carry on as it normally does. In the second experiment (named ASE), observed rainfall fields derived from a mosaic of hourly CMORPH rainfall retrievals were used to force the Noah land model replacing the model-generated precipitation fields, while the initial soil conditions were defined in the same way as in the CTRL experiment. In this experiment, the radar data ingestion was performed for the entire simulation period (96 hours) at each model time step after linearly interpolating the hourly rainfall data.

4.4.2 Results – Physical interpretation

The application of the satellite data ingestion technique and the evaluation methodology followed for the EUROPE/WRF study is similar to the one used for the USA/POSEIDON study (see paragraphs 4.3.3 to 4.3.5). The only difference pertains to the fact that the ASE2 experiment was not performed in the EUROPE/WRF study, as our main goal here was to carry out a sensitivity analysis of the assimilation technique rather than confirm its QPF capability.

The accumulated rainfall fields obtained from the two model experiments (CTRL and ASE) for the 6-hour period ending at 18 UTC on 23 August 2005 are shown in Fig. 4.12 (panels a and b). The third map (Fig. 4.12c) corresponds to the respective rainfall fields as obtained by the CMORPH satellite retrievals. As indicated in Fig. 4.12, after 66 hours of model integration the CTRL run (panel a) produces some of the patches of the 6-hourly

accumulated precipitation fields retrieved by CMORPH (panel c). However, the implementation of ASE experiment (panel b) seems to bring the modelled rainfall fields closer to the truth (i.e., the satellite-retrieved rainfall fields), by enhancing rain rates in the regions where they were underestimated in the CTRL run (e.g., over south Romania) and limiting them in the regions where they were overestimated (e.g., over west Romania). We have to point out though that both runs did not capture the quantity of rainfall over South Romania as indicated by the "dark" precipitation patch in Fig. 4.12a.

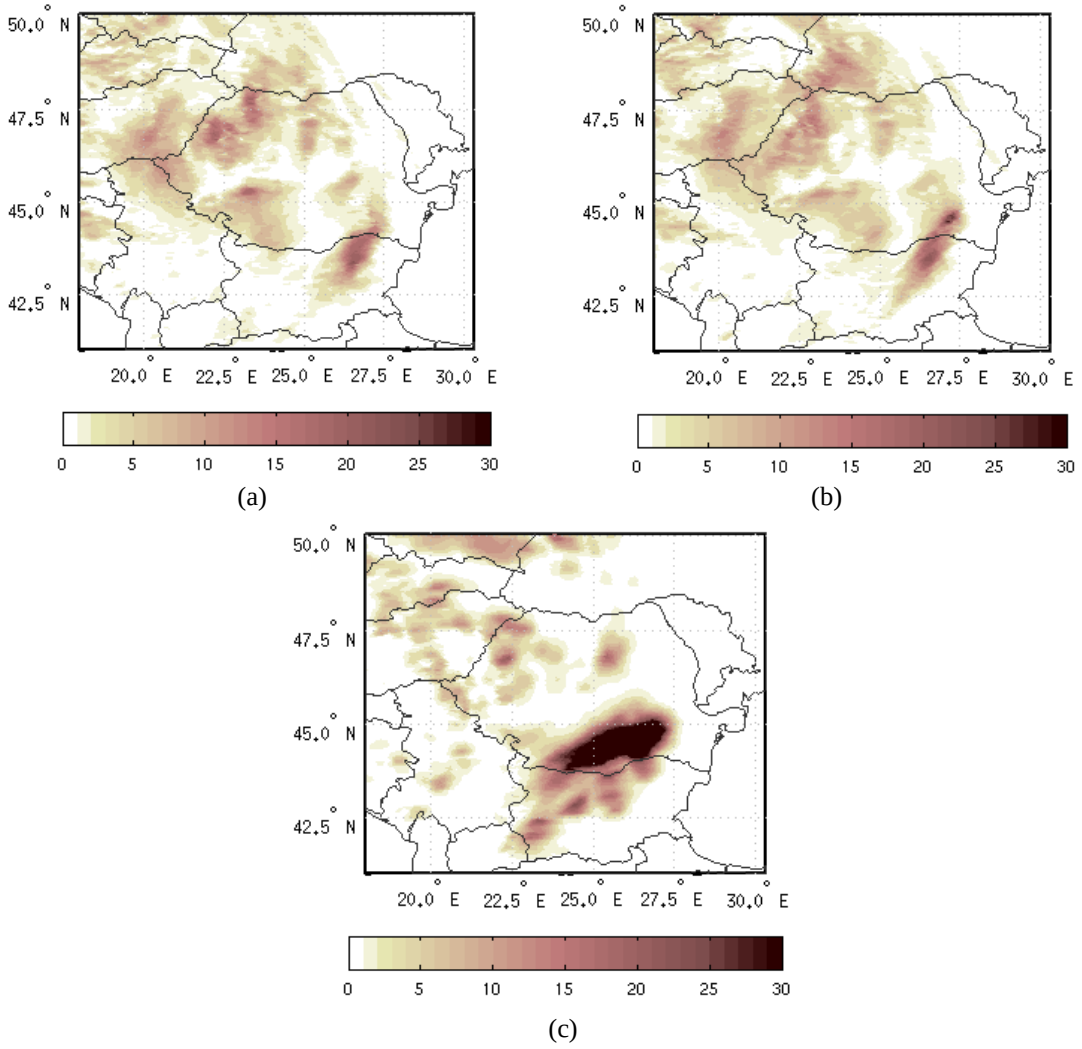
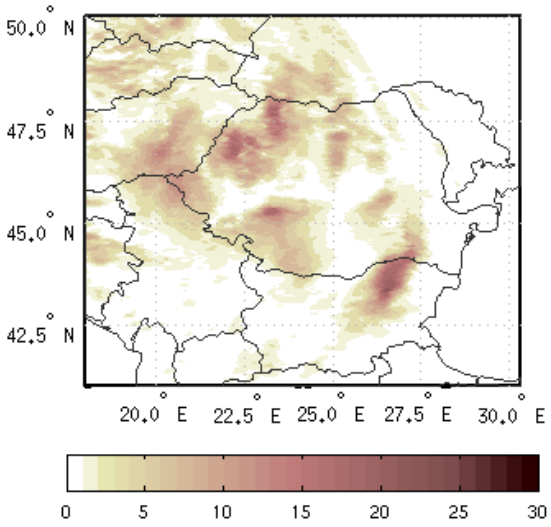
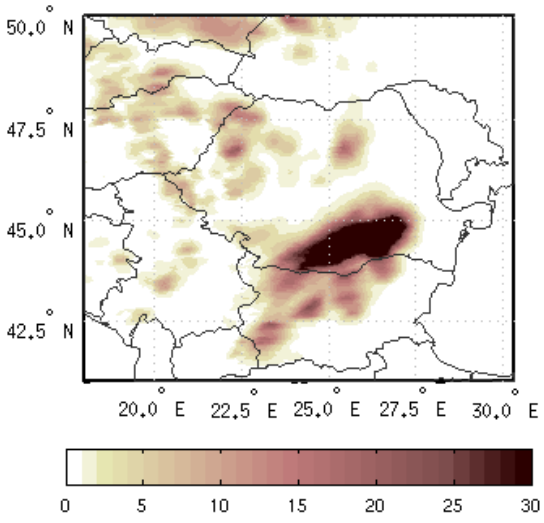


Figure 4.12. Rainfall fields (in mm), as predicted by CTRL (a) and ASE (b) experiments and as retrieved by CMORPH (c), for a 6-h period ending at 18 UTC 23 August 2005.

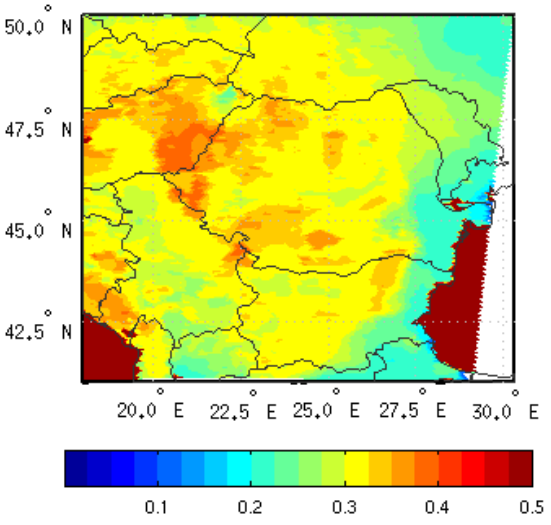
In 4.3.6, as part of the USA/POSEIDON study we discussed the link between surface state and the atmospheric boundary layer (and consequentially precipitation) at the short-term periods resolved by mesoscale models. These feedbacks were the main focus of the EUROPE/WRF case as well. Fig. 4.13 facilitates a graphical investigation of the interactions between rainfall, near surface soil moisture and latent heat flux, on the basis of their differences between CTRL and ASE experiments at 18 UTC on 23 August 2005 for the area confined by 18° E, 31° E, 42° N and 50° N.



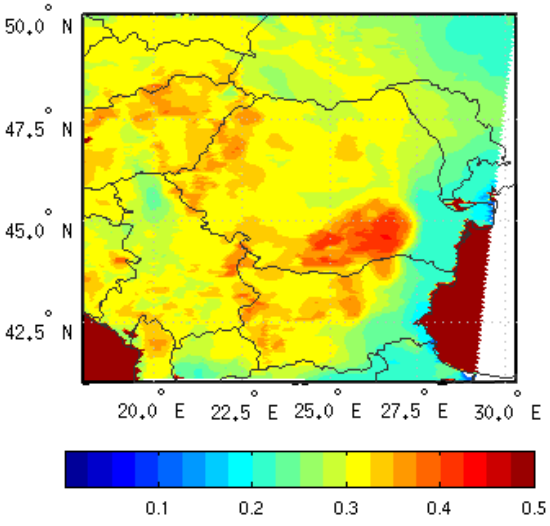
(a) CTRL model predicted rainfall, mm



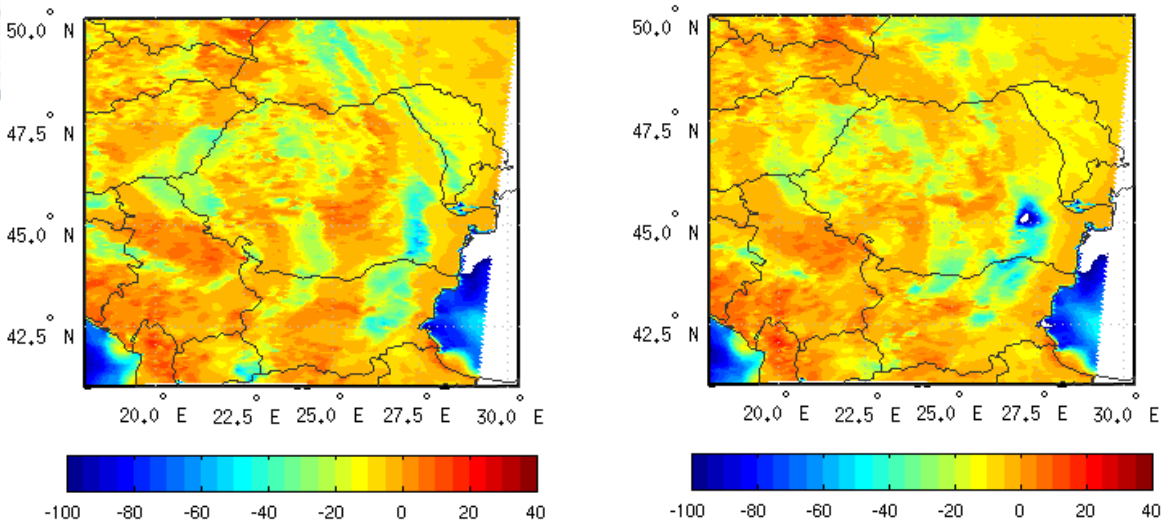
(b) CMORPH rainfall, mm



(c) soil moisture (5 cm), $m^3 m^{-3}$



(d) soil moisture (5 cm), $m^3 m^{-3}$



(e) latent heat flux, W m^{-2} (f) latent heat flux, W m^{-2}

Figure 4.13. Surface fields of land-air interaction properties at 18 UTC 23 August 2005 for the area under investigation in the EUROPE/WRF study. Left and right panels show CTRL and ASE experiments, respectively. The top panels show 6-hour accumulated rainfall (mm) for the period 12-18 UTC on 23 August 2005 as predicted by the model (panel a) and captured by CMORPH (panel b). The middle and bottom panels show model-predicted land surface parameters at 18 UTC, e.g. soil moisture at 5 cm (panels c and d) and latent heat flux (panels e and f). Negative values indicate higher latent heat flux.

The 6-hour (12-18 UTC) accumulation of rainfall fields ingested in the Noah land model (Fig. 4.13b) appear to have a strong impact on the predicted near surface soil moisture content (Fig. 4.13, panel d as opposed to panel c). Soil moisture modulations seem to follow the spatial rainfall patterns ingested in the land model during the ASE experiment. The feedback is especially evident in the areas of south-eastern Romania, where intense rainfall causes significant increase in near surface soil moisture (at the order of 0.05 to 0.15). The modified soil moisture fields appear to further influence the latent heat flux spatial distribution (Fig. 4.6, panel f as opposed to panel e). Latent heat flux is higher (by about 25 to 50 Wm^{-2}) in the areas where soil moisture shows the highest increase from CTRL to ASE experiment, e.g. south-eastern Romania. The soil temperature and sensible heat flux spatial distributions (graphs not shown here) are also affected by the modulation of soil moisture fields, but the sign of these changes is weakly defined. Nevertheless, we note similar temperatures and reduced sensible heat fluxes in areas where precipitation is ingested due to radar data.

Applying the methodology followed in 4.3.6 to quantify the impact of the assimilation technique on the prescribed land-air interaction parameters, the distributions of the relative differences between the ASE and CTRL experiments are plotted in Fig. 4.14, separately for the grid points where the 6-hour CMORPH accumulated rainfall is greater than the corresponding 6-hour rainfall predicted by the model CTRL experiment, and vice versa. Based on Eq. 4.6, the mean values of these relative differences (MRDs) are summarized in Table 4.2 that also provides ratios of MRD values of soil moisture, soil temperature, and fluxes of latent and sensible heat to the MRD of precipitation.

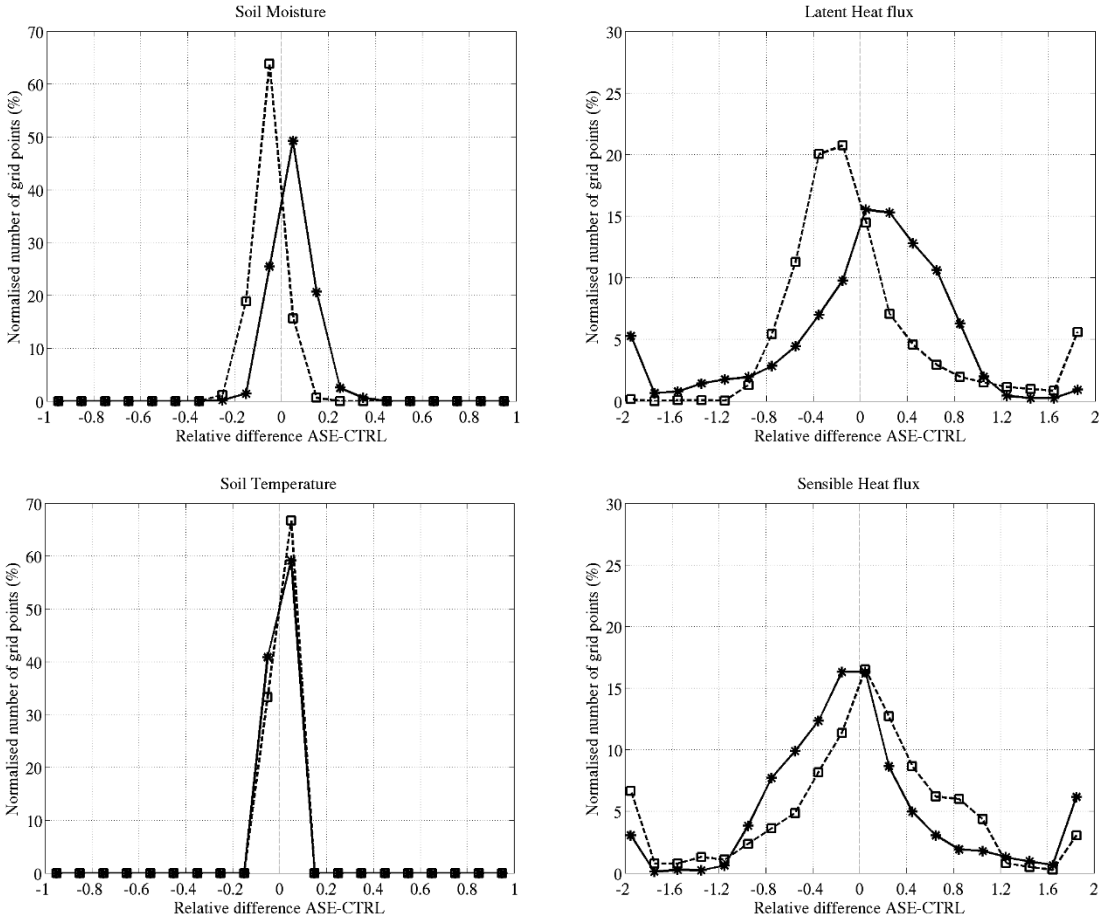


Figure 4.14. Frequency distributions of the relative differences between ASE and CTRL experiments for the model-predicted land-surface parameters: near surface soil moisture ($m^3 m^{-3}$), latent heat flux ($W m^{-2}$), near surface soil temperature (°C) and sensible heat flux ($W m^{-2}$). For each parameter, the relative difference is calculated for the grid points where the 6-h accumulated rainfall as observed by CMORPH is higher than the corresponding rainfall predicted by the model CTRL experiment (stars, solid line), and elsewhere (squares, dashed line).

Table 4.2. Mean relative differences (MRDs) between ASE and CTRL experiments and mean relative differences normalized with the mean relative difference of 6-h accumulated rainfall (MRD/MRD_{rain}) for the parameters shown in **Figure 4.13**.

	CMORPH rainfall > CTRL rainfall		CMORPH rainfall < CTRL rainfall	
	MRD	MRD/MRD_{rain}	MRD	MRD/MRD_{rain}
6-h acc. rainfall (12-18 UTC)	0.75	-	-0.78	-
Soil moisture	0.049	0.065	-0.051	0.066
Latent heat flux	0.184	0.245	-0.15	0.192
Soil temperature	-0.007	-0.01	0.008	-0.01
Sensible heat flux	-0.099	-0.132	0.047	-0.061
Bowen ratio	-0.223	-0.297	0.163	-0.209
1-h acc. rainfall (18-19 UTC)	0.274	0.365	-0.216	0.277

Fig. 4.14 and Table 4.2 verify that the areas, where the precipitation ingested in the land model is greater than the respective rainfall passed onto the land model from the atmospheric model, are characterized by increased values of soil moisture and latent heat flux, with the normalized mean relative difference being about 7% for soil moisture and 24% for latent heat flux. Similar values (-7% and -19%, respectively) are obtained in the case of precipitation ingested in the land model being lower than the respective predicted rainfall from the CTRL experiment. The effect on latent heat fluxes exhibits wider spread and more skewed distributions towards larger values (both positive and negative), caused by the non-linear land surface processes. In terms of soil temperature, the effect is not evident (mean relative difference is below 1%), but in sensible heat fluxes the relative difference is small but not negligible (mean relative differences at the order of 7-13%). Therefore, similar to the USA/POSEIDON case, ingested precipitation has the inverse impact on sensible heat flux than latent heat flux. The combined effect of heat fluxes is a significant decrease (increase) in the Bowen ratio in areas where the CMORPH rainfall ingested to the model is higher (lower) than the corresponding rainfall predicted by the model CTRL experiment, as indicated by the respective frequency distribution shown in

Fig. 4.15 (panel a). The mean relative differences are about -30% and 21% for the two cases, respectively. The net effect on the subsequent model prediction of precipitation (1 hour of accumulation: 18-19 UTC) is consequentially high, as shown in Fig. 4.15 (panel b), with mean increase (decrease) in precipitation in the range of 30% in areas where CMORPH rainfall rates are higher (lower) than CTRL.

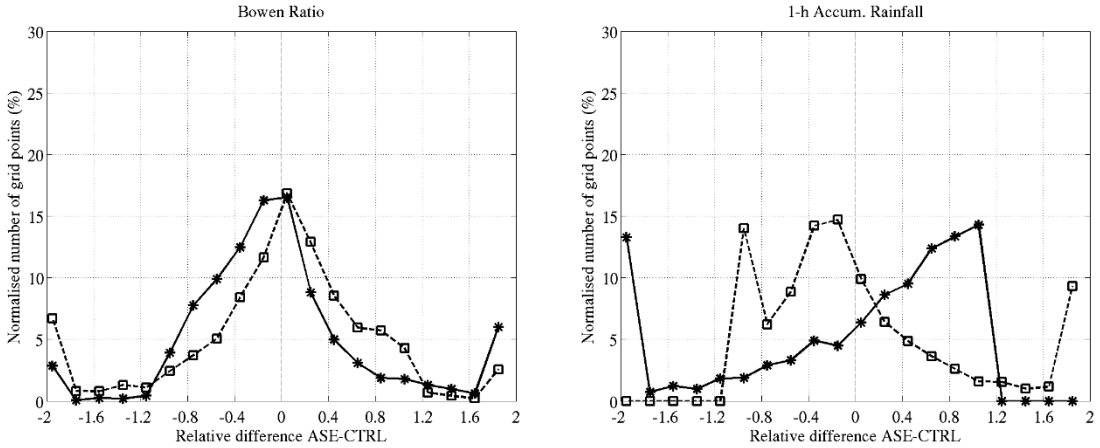


Figure 4.15. As in **Figure 4.14**, but for the Bowen ratio (panel a) and the subsequent accumulated rainfall for the period 18-19 UTC on 23 August 2005 (panel b).

4.5 Comparison – Feedback mechanisms

Comparing the evaluation and physical interpretation results between the two case studies, we observe that in both cases precipitation ingestion from remotely-sensed data (NEXRAD vs CMORPH) in the land surface schemes (OSU vs Noah) of the weather forecasting models (POSEIDON vs WRF), as a replacement of the model-generated precipitation, significantly improves the subsequent simulated precipitation fields through the improvement of the representation of land surface properties and turbulent fluxes. The effect in the USA/POSEIDON study is slightly stronger than the EUROPE/WRF study and this can be attributed to the different domains used in combination with the different model dynamics.

Fig. 4.16 summarizes the feedback mechanisms caused by replacing model predicted rainfall with NEXRAD/CMORPH rainfall. In areas where NEXRAD/CMORPH rainfall accumulation is larger than the CTRL model prediction (i.e., precipitation increase) soil moisture and consequently evaporation rates increase. Increased evaporation would cause an increase of cloud formation and subsequent precipitation. Increased evaporation would moderate the increase in soil moisture through negative feedback, and decrease land surface temperature, which is further affected by the decrease of net radiation caused by increase of cloud formation. Evaporation increase and surface temperature decrease cause increase and decrease of latent and sensible heat fluxes, respectively. The net effect is a decrease of Bowen ratio and increase of subsequent precipitation predicted by the model. The opposite feedback mechanism would take place in a precipitation decrease scenario by replacing CTRL rainfall with NEXRAD/CMORPH rainfall.

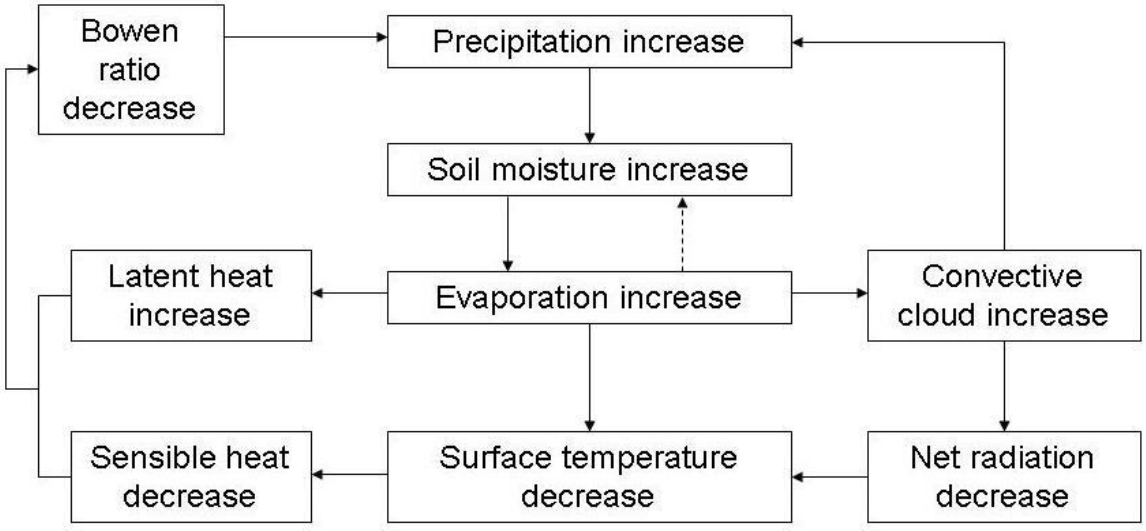


Figure 4.16. Feedback mechanisms. Solid lines stand for positive feedback, and dashed lines stand for negative feedback.

4.6 Discussion and limitations

The coupling procedure between NWP models and LSMs is subject to substantial model biases and errors that may negatively impact the quality of their output due to the

positive feedbacks caused by the non-linear land-atmosphere interactions. This was one of the major motivations for the development of the advanced Land Data Assimilation Systems (LDAS), described thoroughly in chapter 1. These systems require information such as precipitation, downward shortwave and longwave radiation, near surface air temperature, humidity, and wind speed. Precipitation and solar radiation are considered as the most important forcing terms due to their significant impact on the water and energy budgets (Cosgrove et al. 2003a). Therefore, the success of LDAS relies primarily on the accuracy of the estimated precipitation and radiation forcing fields. Careful testing is also necessary with respect to the LSMs included in the LDAS in terms of soil moisture representation (Schaake et al 2004) and initialization procedures (Rodell et al. 2005).

Moreover, some limitations are imposed on the proposed technique by the uncertainties and errors associated with the radar and satellite rainfall estimates (Ciach et al. 2007). The use of CMOPRH data as the precipitation source of the assimilated fields in the EUROPE/WRF study could be considered as such a limitation, due to the overestimation of rainfall by the morphing technique as seen in chapter 2 (par. 2.4.1). The use of CMOPRH data in the study was substantiated by the need to integrate WRF in a domain much bigger than the area influenced by the mesoscale cyclone in combination with the lack of other extensive remotely-sensed rainfall data (e.g., radar) for Europe at that time. In addition, this limitation is lifted with the use of relative differences as a tool to analyse and compare the feedbacks of modelled land-surface properties between the numerical experiments with and without the ingestion of the CMOPRH precipitation data in the WRF land surface scheme.

Chapter 5 Summary and Conclusions

5.1 Summary and major results

The first part of the thesis presented an in-depth investigation of the properties of remotely-sensed rainfall error propagation in the prediction of near-surface soil moisture from a LSM. Specifically, two error sources were contrasted: in rainfall forcing due to estimation error by remote sensing techniques and in the representation of land-atmospheric processes due to LSM uncertainty (in this particular study we used the Community Land Model, version 3.5). CLM3.5 was forced by three remotely-sensed precipitation products, namely, two satellite-based estimates, the NASA-TRMM multi-satellite precipitation analysis and the NOAA Climate Prediction Centre morphing technique, and a rain gauge-adjusted radar-rainfall product from the WSR-88D network. The error analysis was performed for the warm seasons of 2004 and 2006 and facilitated through the use of in-situ measurements of soil moisture, rainfall and other meteorological variables from a network of stations capturing the State of Oklahoma (Oklahoma Mesonet). The study also presented a rigorous benchmarking of the Mesonet network as to its accuracy in deriving area rainfall estimates at the resolution of satellite products (0.25 degree & 3-hourly) through comparisons against the most definitive measurements of a smaller yet denser network of rain gauges in Southwestern Oklahoma (Micronet). The study compared error statistics between modelling and precipitation error sources and between the various remote sensing techniques. Results were contrasted between the relatively moist summer period of 2004 to the drier summer period of 2006 indicating model and error propagation dependencies. An inter-comparison between rainfall and modelling error showed that the two error sources are of similar magnitudes in the case of high modelling accuracy (i.e., 2004), while rainfall forcing error contribution to soil moisture prediction uncertainty can be lower when the model's efficiency skill is relatively low (i.e., 2006).

The second part of the thesis aimed at resolving one other deficiency of LSMs (and in particular CLM3.5). In regional-scale climate applications LSMs are commonly coupled to atmospheric models to close the surface energy, mass and carbon balance. LSMs in

these applications are used to resolve the momentum, heat, water and carbon vertical fluxes, accounting for the effect of vegetation, soil type and other surface parameters, while lack of adequate resolution prevents using them to resolve horizontal sub-grid processes. Specifically, LSMs resolve the large-scale runoff production associated with infiltration excess and sub-grid groundwater convergence, but they neglect the effect from loosing streams to groundwater. Through the analysis of observed data of soil moisture obtained from the Oklahoma Mesoscale Network stations and land surface temperature derived from MODIS, evidence was provided that the regional scale soil moisture and surface temperature patterns are affected by the rivers. This is demonstrated on the basis of CLM3.5 simulations. It was shown that the model cannot reproduce the features of the observed soil moisture and temperature spatial patterns that are related to the underlying mechanism of reinfiltration of river water to groundwater. Therefore, a simple parameterization of this process in CLM3.5 was implemented, showing the ability to reproduce the soil moisture and surface temperature spatial variabilities that relate to the river distribution at regional scale. The CLM3.5 with this new parameterization was used to evaluate impacts of the improved representation of river-groundwater interactions on the simulated water cycle parameters and the surface energy budget at the regional scale.

The results of the first two parts of the thesis showed the significance of using an advanced land surface model with improved parameterizations for investigating the land-atmosphere interactions at the regional scale. Moreover, these results justify the selection of CLM3.5 as the main land surface model in current and future land data assimilation systems. The advanced atmospheric models used as the atmospheric part of these LDASs as well as their influence from land data assimilation techniques was the focus of the third part of the thesis. This part investigated the effect of forcing the land surface schemes of atmospheric mesoscale models with remotely-sensed rainfall data instead of the model-generated rainfall fields. The goal was to provide improved surface conditions for the atmospheric models to achieve accurate simulations of the mesoscale circulations that could significantly affect the timing, distribution and intensity of convective precipitation. The performance of the approach was evaluated in two separate case studies. The USA/POSEIDON study utilized the POSEIDON forecasting system in combination with radar rainfall data from the USA NEXRAD network to explore a mesoscale convective

system that occurred over the US Great Plains in July 2004. The experimental design included multiple runs covering a variety of forcing periods. Continuous data integration was initially used to investigate the sensitivity of the model's performance in varying soil state conditions, while shorter time windows prior to the storm event were utilized to assess the effectiveness of the procedure for improving convective precipitation forecasting. Results indicated that continuous integration of radar rainfall data brings the simulated precipitation fields closer to the observed ones, as compared to the control simulation. The precipitation forecasts (up to 48 h) appeared improved also in the cases of shorter integration periods (24 and 36 h), making this technique potentially useful for operational settings of weather forecasting systems. The second study (i.e., the EUROPE/WRF study) utilized the WRF system in combination with satellite rainfall data from the CMORPH technique to explore a mesoscale convective system that occurred over south-eastern Europe in August 2005. The experimental design was similar to the first case study with respect to the continuous data integration, which facilitated the in-depth investigation of the physical characteristics of modelled land-atmosphere interactions. Both studies showed that the precipitation ingestion from remotely-sensed data (NEXRAD vs CMORPH) in the land surface schemes (OSU vs Noah) of the weather forecasting models (POSEIDON vs WRF), as a replacement of the model-generated precipitation, significantly improves the subsequent simulated precipitation fields through the improvement of the representation of land surface properties and turbulent fluxes. The effect in the USA/POSEIDON study is slightly stronger than the EUROPE/WRF study, which can be attributed to the different domains used in combination with the different model dynamics. The mechanism proposed in Fig 4.16 summarizes the modelled land-atmosphere interaction feedbacks and significantly contributes to the understanding of the complex processes that characterize the water cycle over land.

5.2 Concluding remarks and future work

The dissertation was facilitated by the collection, post-processing and application of hydrometeorological data from ground stations, weather radar and satellites, as well as the application of advanced land surface and meteorological models. These state-of-the-art

tools with the use of new model parameterizations and data assimilation techniques facilitated an in-depth examination of the interactions between groundwater, land surface state (soil moisture and temperature, turbulent fluxes etc.) and atmospheric boundary layer (convection, precipitation, etc.), at the regional scale and for time frames ranging from a few days to a few months.

One of the main achievements of this dissertation was the dynamic correction of numerical weather prediction model prognostic results through the development of a new assimilation technique enabled by the offline coupling between the model land surface schemes and remotely-sensed rainfall data. This part of the study was based on the investigation of specific cases of special meteorological interest, such as convective thunderstorms and severe floods, which occurred over parts of Europe and North America in the recent past.

At the same time, the dissertation contributed to the hydrologic science with the development of a new parameterization scheme of underground water cycle processes in an advanced land surface model and improved our understanding of the complex interactions between land surface streams and groundwater at the regional scale. A detailed study of remotely-sensed rainfall error propagation in the simulated soil moisture fields of the same land surface model was also performed leading to significant conclusions with respect to the model uncertainty characterization and respective climatological effects.

The work contained herein offers significant tools to the scientific community to further explore ways of reducing the disastrous effects of the combination of continental thunderstorms induced by intense cyclonic systems with specific topographic domains and river networks. The study also contributes to current and future work on advancing the land data assimilation systems used for producing global fields of Essential Climate Variables (ECVs), for instance soil moisture and soil temperature, which are deemed necessary for improved representation of global climatic variability.

Abstract

The dissertation focuses on the study of the continental water cycle and combines advanced atmospheric and land surface modelling systems with remotely-sensed data to thoroughly investigate the interactions between water cycle processes taking place underground, at the land surface and in the lower atmospheric boundary layer. The first part of the dissertation includes the study of the propagation of error of remotely-sensed precipitation data to the soil moisture fields simulated by an advanced land surface model (LSM). The study facilitates the in-depth investigation and comparison of two major error sources in the simulation of land surface state properties at the regional scale, namely the error induced by rainfall forcing and the error induced by model internal parameterizations. Through this error sensitivity analysis, two other major contributions are achieved; namely, the comparison of the performance of three extensive sources of remotely-sensed rainfall data (NEXRAD, TRMM and CMORPH) as well as the evaluation of the performance of the advanced LSM (Community Land Model, version 3.5). The second part of the dissertation investigates the interaction between water reservoirs over land (rivers, lakes etc.) and the underlying groundwater at the regional/seasonal scale. The study improves the CLM3.5 parameterization of the interactions between surface- and groundwater (TOPMODEL approach) by considering the discharge of land surface streams into groundwater that was not accounted before. The results show significant improvement in the simulations of soil moisture and other land surface properties in the areas including or surrounding major river networks. Both studies identify similar climatological effects of the prescribed analyses. Specifically, the CLM error statistics and the new river-groundwater interaction parameterization highly depend on the wetness conditions that prevail during the simulation period (improved results on wet areas or seasons). The third part of the dissertation discusses the development and application of a new assimilation technique through the ingestion of remotely-sensed precipitation data in the land surface schemes of advanced mesoscale models. The technique significantly improves the models' quantitative precipitation forecasting capability in cases of extreme thunderstorms, thus offering a valuable tool to meteorologists worldwide. The assimilation technique is tested for two different advanced weather forecasting systems (i.e., the POSEIDON and the WRF

systems) in two different continental regimes (i.e., continental USA and Europe, respectively) with two different sources of remotely-sensed rainfall for the LSM data ingestion (i.e., NEXRAD and CMORPH estimates, respectively). Moreover, an extensive feedback investigation is performed, advancing our understanding of the complex land-atmosphere interaction processes.

Περίληψη

Η διατριβή επικεντρώνεται στη μελέτη του ηπειρωτικού υδάτινου κύκλου και συνδυάζει προηγμένα εδαφικά και ατμοσφαιρικά μοντέλα με δεδομένα τηλεπισκόπισης για τη διερεύνηση των αλληλεπιδράσεων μεταξύ των διεργασιών του υδάτινου κύκλου που πραγματοποιούνται κάτω από την επιφάνεια της γης, στην επιφάνεια της αλλά και στο ατμοσφαιρικό οριακό στρώμα. Το πρώτο μέρος της διατριβής περιλαμβάνει τη μελέτη της διάδοσης του σφάλματος που χαρακτηρίζει τα δεδομένα βροχόπτωσης από τηλεπισκόπιση στα πεδία εδαφικής υγρασίας που προσομοιώνονται από ένα προηγμένο εδαφικό μοντέλο. Η μελέτη διευκολύνει την εις βάθος διερεύνηση και σύγκριση δύο σημαντικών πηγών σφάλματος στη προσομοίωση των ιδιοτήτων της επιφανειακής εδαφικής κατάστασης σε περιοχική κλίμακα, δηλαδή το σφάλμα που οφείλεται στις εκτιμήσεις βροχόπτωσης και το σφάλμα που προκαλείται από τις εσωτερικές παραμετροποιήσεις του μοντέλου. Μέσω αυτής της ανάλυσης ευαισθησίας, επιτυγχάνεται επίσης η σύγκριση των επιδόσεων τριών σημαντικών συνόλων δεδομένων βροχόπτωσης από τηλεπισκόπιση (NEXRAD, TRMM και CMORPH) καθώς και η αξιολόγηση των επιδόσεων του προηγμένου εδαφικού μοντέλου (CLM3.5). Το δεύτερο μέρος της διατριβής διερευνά την αλληλεπίδραση μεταξύ των επιφανειακών υδάτων (ποτάμια, λίμνες κ.λπ.) και των υποκείμενων υπόγειων υδάτων, επίσης σε περιοχική και εποχική κλίμακα. Η μελέτη βελτιώνει την παραμετροποίηση του CLM3.5 ως προς τις αλληλεπιδράσεις μεταξύ επιφανειακών και υπόγειων υδάτων (προσέγγιση TOPMODEL), λαμβάνοντας υπόψη τη διήθηση των επιφανειακών υδάτων στα υπόγεια ύδατα που δεν υπήρχε μέχρι τώρα. Τα αποτελέσματα δείχνουν σημαντική βελτίωση στις προσομοιώσεις των πεδίων εδαφικής υγρασίας και άλλων επιφανειακών εδαφικών παραμέτρων στις περιοχές που περιλαμβάνουν μεγάλους ποταμούς ή γύρω από αυτούς. Και οι δύο προαναφερόμενες αναλύσεις παρουσιάζουν παρεμφερή κλιματολογικά χαρακτηριστικά. Συγκεκριμένα, τόσο οι στατιστικές των σφαλμάτων στο CLM όσο και η νέα παραμετροποίηση αλληλεπίδρασης μεταξύ ποταμών και υπόγειων υδάτων εξαρτώνται σε μεγάλο βαθμό από τις συνθήκες υγρασίας που επικρατούν κατά την περίοδο των προσομοιώσεων (καλύτερα αποτελέσματα στις περιπτώσεις υγρών περιοχών και περιόδων). Το τρίτο μέρος της διατριβής πραγματεύεται την ανάπτυξη και εφαρμογή μιας νέας τεχνικής αφομοίωσης δεδομένων βροχόπτωσης από τηλεπισκόπιση στα εδαφικά

σχήματα προηγμένων μετεωρολογικών μοντέλων. Η τεχνική βελτιώνει σημαντικά την ικανότητα πρόγνωσης βροχοπτώσεων σε περιπτώσεις έντονων καταιγιδοφόρων συστημάτων, προσφέροντας έτσι ένα πολύτιμο εργαλείο στους μετεωρολόγους παγκοσμίως. Η τεχνική αφομοίωσης αξιολογείται για δύο διαφορετικά προηγμένα συστήματα πρόγνωσης καιρού (συστήματα POSEIDON και WRF) σε δύο διαφορετικές ηπειρωτικές περιοχές (ΗΠΑ και Ευρώπη, αντίστοιχα) με χρήση δύο διαφορετικών πηγών δεδομένων βροχόπτωσης για την εισαγωγή στο εδαφικό σχήμα (NEXRAD και CMORPH, αντίστοιχα). Επιπλέον, πραγματοποιείται ενδελεχής μελέτη για την καλύτερη κατανόηση των σύνθετων διεργασιών αλληλεπίδρασης και ανατροφοδοτήσεων μεταξύ εδάφους και ατμόσφαιρας.

Bibliography

Albergel, C., C. Rüdiger, T. Pellarin, J.-C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Piguet, and E. Martin, 2008: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. Hydrol. Earth Syst. Sci., 12, 1323-1337.

Alley, W. M., Healy, R. W., LaBaugh, J. W., and Reilly, T. E., 2002: Flow and storage in groundwater systems. Science, v. 296, p. 1985-1990.

Anagnostou, E. N., V. Maggioni, E. I. Nikolopoulos, T. Meskele, F. Hossain and A. Papadopoulos, 2010: Benchmarking High-Resolution Global Satellite Rainfall Products to Radar and Rain-Gauge Rainfall Estimates. IEEE Trans. Geosci. Remote Sens., 48, 1667-1683.

Anyah, R. O., C. P. Weaver, G. Miguez-Macho, Y. Fan, and A. Robock, 2008: Incorporating water table dynamics in climate modeling: 3. Simulated groundwater influence on coupled land-atmosphere variability. J. Geophys. Res., 113, D07103, doi:10.1029/ 2007JD009087.

Barnston, A. G., 1992: Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score. Wea. Forecasting, 7, 699-709.

Beljaars, A. C. M., P. Viterbo, M. J. Miller, and A. K. Betts, 1996: The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil- moisture anomalies. Mon. Wea. Rev., 124, 362–383.

Betts, A. K., 1986: A new convective adjustment scheme. Part E. Observational and theoretical basis. Quart. J. Roy. Meteor. Soc., 112, 677–691.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and Arctic air mass data sets. Quart. J. Roy. Meteor. Soc., 112, 693–709.

Betts, A. K., 2004: Understanding hydrometeorology using global models. Bull. Amer. Meteor. Soc., 85, 1673-1688.

Beven, K. J., and M. J. Kirkby, 1979: A physically-based variable contributing area model of basin Hydrology. Hydrol. Sci. Bull., 24, 43–69.

Bierkens, M. P. and B. J. J. M. van den Hurk, 2007: Groundwater convergence as a possible mechanism for multi-year persistence in rainfall. Geophys. Res. Lett., 34, doi:10.1029/2006GL028396.

Blankenship C. B., J. L. Case, B. T. Zavodsky, and W. L. Crosson, 2016: Assimilation of SMOS Retrievals in the Land Information System. IEEE Trans. Geosci. Remote Sens., 54, 6320-6332.

Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson, and Z.-L. Yang, 2002: The land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model. J. Climate, 15, 3123-3149.

Branstetter, M., 2001: Development of a Parallel River Transport Algorithm and Applications to Climate Studies. PhD thesis, University of Texas at Austin. J. Famiglietti, supervisor.

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson and M. D. Eilts, 1995: The Oklahoma Mesonet: A Technical Overview. J. Atmos. Ocean. Technol., 12, 5-19.

Buckingham, E., 1914: On physically similar systems; illustrations of the use of dimensional equations. Phys. Rev., 4, 345–376. doi:10.1103/PhysRev.4.345.

Budyko, M. I., 1974: Climate and Life. Academic Press, 508 pp.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts, 1996: Modeling of land surface evaporation by four schemes and comparison with FIFE observations. J. Geophys. Res., 101, 7251–7277.

Chen, F. Z., Z. I. Janjic, and K. Mitchell, 1997: Impact of atmospheric surface-layer parameterization in the new land-surface scheme of the NCEP mesoscale Eta model. Boundary Layer Meteorol., 48.

Chen, J., and P. Kumar, 2001: Topographic influence on the seasonal and interannual variation of water and energy balance of basins in North America. J. Climate, 14(9), 1989–2014.

Cherkauer K. A., and D. P. Lettenmaier, 1999: Hydrologic effects of frozen soils in the upper Mississippi River basin. J. Geophys. Res. Atmos., 104 (D16), 19599-19610.

Ciach, G. J., W. F. Krajewski, and G. Villarini, 2007: Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. J. Hydrometeor., 8, 1325-1347.

Clark, C. A., and R. W. Arritt, 1995: Numerical simulations of the effect of soil moisture and vegetation cover on the development of deep convection. J. Appl. Meteorol., 34, 2029–2045.

Cohen, D., M. A. Person, R. Daannen, S. Locke, D. Dahlstromn, V. Zabielski, T. C. Winter, D. O. Rosenburry, H. Wright, E. Ito, J. L. Nieber, and W. J. Gutowski Jr., 2006: Groundwater-supported evapotranspiration within glaciated watersheds and under conditions of climate change. J. Hydrol., 320, 484–500.

Cosgrove, B. A., D. Lohmann, K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, C. Marshall, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley, and J. Meng, 2003a: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. J. Geophys. Res., 108(D22), 8842, doi:10.1029/2002JD003118.

Cosgrove, B. A., D. Lohmann, K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, and J. D. Tarpley, 2003b: Land surface model spin-up behaviour in the North American Land

Data Assimilation System (NLDAS), J. Geophys. Res., 108, D228845, doi:10.1029/2002JD003316.

Dahm, C. N., M. B. Baker, D. I. Moore, and J. R. Thibault, 2003: Coupled biogeochemical and hydrological responses of streams and rivers to drought. Freshwater Biology, 48, 1219–1231.

David, C. H., D. J. Gochis, D. R. Maidment, W. Yu, D. N. Yates, and Z. -L. Yang, 2009: Using NHDPlus as the Land Base for the Noah-distributed Model. Trans. GIS, 13(4), 363–377.

De Jeu, R., and W. Dorigo, 2016: On the importance of satellite observed soil moisture, Int. J. Appl. Earth Observ. Geoinf., 45 (Part B), 107–109.

Dorigo W., and R. De Jeu, 2016: Satellite soil moisture for advancing our understanding of earth system processes and climate change, Int. J. Appl. Earth Observ. Geoinf., 48, 1–4.

Drusch, M., and P. Viterbo, 2007: Assimilation of Screen-Level Variables in ECMWF's Integrated Forecast System: A Study on the Impact on the Forecast Quality and Analyzed Soil Moisture. Mon. Wea. Rev., 135, 300–314.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. J. Geophys. Res., 105(D20), 24,823–24,838.

Ebert, E. E., J. E. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bull. Amer. Meteor. Soc., 88, 47–64.

Entekhabi, D., 1995: Recent advances in land-atmosphere interaction research. U.S. Natl. Rep. Int. Union Geod. Geophys. 1991– 1994, Rev. Geophys., 33, 995– 1003.

Entekhabi, D., E. G. Njoku, P. Houser, M. Spencer, T. Doiron, Y. Kim, J. Smith, R. Girard, S. Belair, W. Crow, T. J. Jackson, Y. H. Kerr, J. S. Kimball, R. Koster, K. C. McDonald, P. E. O'Neill, T. Pultz, S. W. Running, J. Shi, E. Wood, and J. Van Zyl, 2004: The Hydrosphere State (Hydros) Satellite Mission: an earth system pathfinder for global mapping of soil moisture and land freeze/thaw. IEEE Trans. Geosci. Remote Sens., 42(10), 2184–2195.

Entekhabi, D., E. G. Njoku, P. E. O'Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, J. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, and J. Van Zyl, 2010: The Soil Moisture Active Passive (SMAP) mission. Proc. IEEE, 98, 704-716.

Entekhabi, D., S. Yueh, P.E. O'Neill, K. Kellogg, A. Allen, R. Bindlish, M. Brown, S. Chan, A. Colliander, W. T. Crow, N. Das, G. De Lannoy, R. S. Dunbar, W. N. Edelstein, J. K. Entin, V. Escobar, S. D. Goodman, T. J. Jackson, B. Jai, J. Johnson, E. Kim, S. Kim, J. Kimball, R. D. Koster, A. Leon, K. C. McDonald, M. Moghaddam, P. Mohammed, S. Moran, E. G. Njoku, J. R. Piepmeier, R. Reichle, F. Rogez, J.-C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, J. Van Zyl, B. Weiss, and R. West 2014. SMAP handbook: Mapping soil moisture and freeze/thaw from space. Publ. JPL 400-1567. NASA, Jet Propulsion Lab., Pasadena, CA.

Famiglietti, J. S., and E. F. Wood, 1994: Multiscale modeling of spatially variable water and energy balance processes. Water Resour. Res., 30, 3061–3078.

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

Fang L., C.R. Hain, X. Zhan, and M.C. Anderson, 2016: An inter-comparison of soil moisture data products from satellite remote sensing and a land surface model. Int. J. Appl. Earth Observ. Geoinf., 48, 34-47.

Fels, S. B., and M. D. Schwarzkopf, 1975: The simplified exchange approximation: A new method for radiative transfer calculations. J. Atmos. Sci., 32, 1475–1488.

Ferrier, B. S., Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. Preprints, 15th Conference on Numerical Weather Prediction, San Antonio, TX, Amer. Meteorol. Soc., 280-283.

Fischer, E.M., S.I. Seneviratne, P.L. Vidale, D. Lüthi, and C. Schär, 2007: Soil Moisture–Atmosphere Interactions during the 2003 European Summer Heat Wave. J. Climate, 20, 5081–5099.

Fulton, R., J. Breidenbach, D-J Seo, and D. Miller, 1998: The WSR-88D Rainfall Algorithm. Wea. Forecasting, 13, 377-395.

Gedney, N., and P. M. Cox, 2003: The sensitivity of global climate model simulations to the representation of soil moisture heterogeneity. J. Hydrometeor., 4, 1265–1275.

Gehne, M., T.M. Hamill, G.N. Kiladis, and K.E. Trenberth, 2016: Comparison of Global Precipitation Estimates across a Range of Temporal and Spatial Scales. J. Climate, 29, 7773–7795, https://doi.org/10.1175/JCLI-D-15-0618.1.

Gochis D. J., and F. Chen, 2003: Hydrological Enhancements to the Community Noah Land Surface Model. Web document, http://www.ucar.edu/library/collections/technotes/technotes.jsp

Gottschalck, J., J. Meng, M. Rodell, and P. Houser, 2005: Analysis of multiple precipitation products and preliminary assessment of their impact on Global Land Data Assimilation System land surface states. J. Hydrometeor., 6, 573–598.

Gutowski, W. J., Jr., C. J. Vörösmarty, M. Person, Z. Ötles, B. Fekete, and J. York, 2002: A coupled land–atmosphere simulation program (CLASP): Calibration and validation. J. Geophys. Res., 107, 4283, doi:10.1029/2001JD000392.

Heidke, P., 1926: Berechnung des erfolges und der gute der windstarkevorhersagen im sturmwarnungsdienst. Georg. Ann., 8, 310-349.

Hossain, F., and E. N. Anagnostou, 2004: Assessment of current passive microwave and infrared based satellite rainfall remote sensing for flood prediction, J. Geophys. Res., 109, D07102, doi:10.1029/2003JD003986.

Hossain, F., and E. N. Anagnostou, 2005a: Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface model parameters on simulation of soil moisture. Adv. Water Resour., 28, 1336-1350.

Hossain, F., and E. N. Anagnostou, 2005b: Using a multi-dimensional satellite rainfall error model to characterize uncertainty in soil moisture fields simulated by an offline land surface model. Geophys. Res. Lett., 32, L15402, doi:10.1029/2005GL023122.

Hossain, F., and E. N. Anagnostou, 2006: A two-dimensional satellite rainfall error model, IEEE Trans. Geosci. Remote Sens, 44, 1511-1522.

Hossain, F., and G.J. Huffman, 2008: Investigating Error Metrics for Satellite Rainfall Data at Hydrologically Relevant Scales. J. Hydrometeor., 9, 563–575.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined sensor precipitation estimates at fine scales. J. Hydrometeor., 8, 38–55.

Jacob R., J. Larson, and E. Ong, 2005: MxN Communication and Parallel Interpolation in CCSM3 Using the Model Coupling Toolkit. Int. J. High Perf. Comp. Appl., 19(3), 293-307.

Janjic, Z. I., 1984: Non-linear advection schemes and energy cascade on semi-staggered grids. Mon. Wea. Rev., 112, 1234–1245.

Janjic, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer and turbulence closure schemes. Mon. Wea. Rev., 122, 927-945.

Janjic, Z. I., 1996a: The Mellor-Yamada level 2.5 scheme in the NCEP Eta Model. 11th Conference on Numerical Weather Prediction, Norfolk, VA, Amer. Meteorol. Soc., 333-334.

Janjic, Z. I., 1996b: The Surface Layer in the NCEP Eta Model. 11th Conf. on Numerical Weather Prediction, Norfolk, VA, Amer. Meteorol. Soc., 354–355.

Janjic, Z. I., 2000: Comments on "Development and Evaluation of a Convection Scheme for Use in Climate Models". J. Atmos. Sci., 57, p. 3686.

Janjic, Z. I., 2002: Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, NCEP Office Note, No. 437, 61 pp.

Janjic, Z.I., 2003: A nonhydrostatic model based on a new approach. Meteorol. Atmos. Phys., 82, 271-285.

Jiang X., G. Y. Niu, and Z. L. Yang, 2009: Impacts of vegetation and groundwater dynamics on warm season precipitation over the Central United States. J. of Geophys. Res. Atmos., 114, D06109.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeor., 5, 487–503.

Ju W. M., J. M. Chen, T. A. Black, A. G. Barr, H. McCaughey, and N. T. Roulet, 2006: Hydrological effects on carbon cycles of Canada's forests and wetlands. Tellus B, 58(1) 16-30.

Kain, J. S., and J. M. Fritsch, 1992: The role of the convective ‘‘trigger function’’ in numerical forecasts of mesoscale convective systems. Meteor. Atmos. Phys., 49, 93–106.

Kallos, G., S. Nickelic, A. Papadopoulos, D. Jovic, O. Kakaliagou, N. Misirlis, L. Boukas, N. Mimikou, G. Sakellaridis, J. Papageorgiou, E. Anadranistakis, and M. Manousakis, 1997: The regional weather forecasting system SKIRON: An overview. Proc. Symp. on Regional Weather Prediction on Parallel Computer Environments, Athens, Greece, University of Athens, 109–122.

Kerr Y. H., P. Waldteufel, J. P. Wigneron, J. M. Martinuzzi, J. Font, and M. Berger, 2001: Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) Mission. IEEE Trans. Geosci. Remote Sens., 39(8), 1729–1735.

Kingston D. G., D. M. Hannah, D. M. Lawler, and G. R. McGregor, 2009: Climate-river flow relationships across montane and lowland environments in northern Europe. Hydrol. Processes, 23 (7), 985-996.

Kirschbaum, D., G. Huffman, G. Skofronick-Jackson, S. Braun, E. Stocker, K. Garrett, E. Jones, R. Adler, H. Wu, A. McNally, and B. Zaitchik, 2017: NASA’s Remotely-sensed Precipitation: A Reservoir for Applications Users. Bull. Amer. Meteor. Soc., doi:10.1175/BAMS-D-15-00296.1.

Kollet, S. J., and R.M. Maxwell, 2006: Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. Adv. Water Resour., 29, 945–958.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. J. Geophys. Res., 105(D20), 24,809–24,822.

Koster, R. D., and M. J. Suarez, 2001: Soil moisture memory in climate models, J. Hydrometeor., 6, 558– 570.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, M. Kistler, A. Berg, R. Reichle, M. Rodell, and J. Famiglietti, 2004a: Realistic initialization of land surface states: impacts on subseasonal forecast skill, J. Hydrometeor., 5, 1049-1063.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchel, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada, 2004b: Regions of strong coupling between soil moisture and precipitation. Science, 305, 1138-1140.

Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E. F. Wood, and J. Sheffield, 2006: Land Information System - An interoperable framework for high resolution land surface modeling, Environ. Model. Softw., 21, 1402-1415.

Kumar, S. V., R. H. Reichle, R. D. Koster, W. T. Crow, and C. D. Peters-Lidard, 2009: Role of subsurface physics in the assimilation of surface soil moisture observations. J. of Hydrometeor., 10, 1534-1547.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the earth's atmosphere. J. Atmos. Sci., 31, 118–133.

Larson J., R. Jacob, and E. Ong, 2005: The Model Coupling Toolkit: A New Fortran90 Toolkit for Building Multiphysics Parallel Coupled Models. Int. J. High Perf. Comp. Appl., 19(3), 277-292.

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

Leese, J., T. Jackson, A. Pitman, and P. Dirmeyer, 2001: meeting summary – GEWEX/BAHC international workshop on soil moisture monitoring, analysis, and prediction for hydrometeorological and hydroclimatological applications. Bull. Amer. Meteor. Soc., 82, 1423–1430.

Liang, X., Z. Xie, and M. Huang, 2003: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model. J. Geophys. Res., 108, 8613, doi:10.1029/2002JD003090.

Lin, L.-F., A. M. Ebtehaj, J. Wang, and R. L. Bras, 2017: Soil moisture background error covariance and data assimilation in a coupled land-atmosphere model. Water Resour. Res., 53, 1309–1335, doi:10.1002/2015WR017548.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. Preprints. 19th Conf. on Hydrology, San Diego, CA, Amer. Meteor. Soc., 1.2.

Liu, Q., R.H. Reichle, R. Bindlish, M.H. Cosh, W.T. Crow, R. de Jeu, G.J. De Lannoy, G.J. Huffman, and T.J. Jackson, 2011a: The Contributions of Precipitation and Soil Moisture Observations to the Skill of Soil Moisture Estimates in a Land Data Assimilation System. J. Hydrometeor., 12, 750–765, https://doi.org/10.1175/JHM-D-10-05000.1

Liu Y.Y., R.M. Parinussa, W.A. Dorigo, R.A.M. De Jeu, W. Wagner, A.I.J.M. Van Dijk, M.F. McCabe, and J.P. Evans, 2011b: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrol. Earth Syst. Sci., 15, 425-436.

Liu Y.Y., W.A. Dorigo, R.M. Parinussa, R.A.M. De Jeu, W. Wagner, M.F. McCabe, J.P. Evans, and A.I.J.M. Van Dijk, 2012: Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sens. Environ., 123, 280-297.

Maggioni, V., R. H. Reichle, and E. N. Anagnostou, 2011: The effect of satellite rainfall error modeling on soil moisture prediction uncertainty. J. Hydrometeor., 12, 413–428.

Maggioni, V., R.H. Reichle, and E.N. Anagnostou, 2012: The Impact of Rainfall Error Characterization on the Estimation of Soil Moisture Fields in a Land Data

Assimilation System. J. Hydrometeor., 13, 1107–1118, https://doi.org/10.1175/JHM-D-11-0115.1

Maggioni V. and P. R. Houser, 2017: Soil Moisture Data Assimilation. In: Park S., Xu L. (eds) Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, Vol. III, Springer, Cham

Mahfouf, J.-F., 1991: Analysis of soil moisture from near-surface parameters. A feasibility study. J. Appl. Meteor., 30, 1534–1547.

Maxwell, R. M., and N. L. Miller, 2005: Development of a coupled land surface and groundwater model. J. Hydrometeor., 6, 233– 247.

McNally, A., S. Shukla, K.R. Arsenault, S. Wang, C.D. Peters-Lidard, and J.P Verdin, 2016: Evaluating ESA CCI soil moisture in East Africa. Int. J. Appl. Earth Obs. Geoinf., doi:10.1016/j.jag.2016.01.001.

Mecklenburg, S., M. Drusch, Y. H. Kerr, M. Martin-Neira, D. Steven, G. Buenadicha, N. Reul, E. Daganzo, R. Oliva, and R. Crapolicchio, 2012: ESA's Soil Moisture and Ocean Salinity Mission: Mission Performance and Operations. IEEE Trans. Geosci. Rem. Sens., 50, no. 5, part 1, 1354–1366.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. Rev. Geophys. Space Phys., 20, 851-875.

Mesinger, F., Z. I. Janjic, S. Nickovic, D. Gavrilov, and D. G. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of an Appalachian redevelopment. Mon. Wea. Rev., 116, 1493–1518.

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

Mitchell, K. E., D. Lohmann, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley, D. P. Lettenmaier, C. H. Marshall, J. K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B. H. Ramsay, and A. A. Bailey, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. J. Geophys. Res., 109, D07S90, doi:10.1029/2003JD003823.

Niu, G. Y., and Z. L. Yang, 2003: The versatile integrator of surface atmospheric processes—Part 2: Evaluation of three topography-based runoff schemes. Global Planet. Change, 38(1–2), 191–208.

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

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

O'Callaghan, J. F., and D. M. Mark, 1984: The extraction of drainage networks from digital elevation data, Comp. Vision Graph. Image Process., 28(3), 323-344.

Oleson, K. W., Y. Dai, G. Bonan, M. Bosilovich, R. Dickinson, P. Dirmeyer, F. Hoffman, P. Houser, S. Levis, G.-Y. Niu, P. Thornton, M. Vertenstein, Z.-L. Yang, and X. Xeng, 2004: Technical description of the Community Land Model (CLM). NCAR Tech. Note NCAR/TN-4611STR, 173 pp.

Oleson, K. W., G.-Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, R. Stöckli, R. E. Dickinson, G. B. Bonan, S. Levis, A. Dai, and T. Qian, 2008: Improvements to the Community Land Model and their impact on the hydrological cycle, J. Geophys. Res., 113, G01021, doi:10.1029/2007JG000563.

Osman Y. R., and M. P. Bruen, 2002: Modelling stream–aquifer seepage in an alluvial aquifer: an improved loosing-stream package for MODFLOW. J. Hydrol., 264, 69–86.

Papadopoulos, A., G. Kallos, S. Nickovic, D. Jovic, M. Dacic, and P. Katsafados, 1997: Sensitivity studies of the surface and radiation parameterization schemes of the SKIRON system. Proc. Int. Symp. on Regional Weather Prediction on Parallel Computer Environments, 15-17 October 1997, Athens, Greece, 155-164.

Papadopoulos, A., G. Kallos, P. Katsafados, and S. Nickovic, 2002: The Poseidon weather forecasting system: An overview. GAOS, 8, 219-237.

Papadopoulos, A., E. Serpetzoglou, and E. N. Anagnostou, 2008: Improving NWP through radar rainfall-driven land surface parameters: A case study on convective precipitation forecasting. Adv. Water Resour., 31, Special Issue on Hydrologic Remote Sensing, 1456-1469.

Papadopoulos, A., E. Serpetzoglou, and E. N. Anagnostou, 2009: Evaluating the impact of lightning data assimilation on mesoscale model simulations of a flash flood inducing storm. Atm. Res., 94 (4), 715–725.

Ramos da Silva, R., and R. Avissar, 2006: The Hydrometeorology of a Deforested Region of the Amazon Basin. J. Hydrometeor., 7, 1028–1042.

Reichle, R.H., R.D. Koster, J. Dong, and A.A. Berg, 2004: Global Soil Moisture from Satellite Observations, Land Surface Models, and Ground Data: Implications for Data Assimilation. J. Hydrometeor., 5, 430–442.

Reichle, R. H., R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). J. Geophys. Res., 112, D09108, doi:10.1029/2006JD008033.

Reichle, R. H., W. T. Crow, R. D. Koster, H. Sharif, and S. P. P. Mahanama, 2008: Contribution of soil moisture retrievals to land data assimilation products. Geophys. Res. Lett., 35, dx.doi.org/10.1029/2007GL031986.

Reichle, R. H., G. J. M. De Lannoy, B. A. Forman, C. S. Draper, and Q. Liu, 2013: Connecting Satellite Observations with Water Cycle Variables through Land Data Assimilation: Examples Using the NASA GEOS-5 LDAS. Surv. Geophys., 35, 577–606, dx.doi.org/10.1007/s10712-013-9220-8.

Robock, A., K.Y. Vinnikov, G. Srinivasan, J.K. Entin, S.E. Hollinger, N.A. Speranskaya, S. Liu, and A. Namkhai, 2000: The Global Soil Moisture Data Bank. Bull. Amer. Meteor. Soc., 81, 1281–1299.

Robock, A., L. Luo, E. F. Wood, F. Wen, K. E. Mitchell, P. R. Houser, J. C. Schaake, D. Lohmann, B. Cosgrove, J. Sheffield, Q. Duan, R. W. Higgins, R. T. Pinker, J. D. Tarpley, J. B. Basara, and K. C. Crawford, 2003: Evaluation of the North American Land Data Assimilation System over the Southern Great Plains during the warm season, J. Geophys. Res., 108, 8846, doi:10.1029/2002JD003245, 2003.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, 2004: The global land data assimilation system. Bull. Amer. Meteor. Soc., 85, 381–394.

Rodell, M., P. R. Houser, A. A. Berg, and J. S. Famiglietti, 2005: Evaluation of ten methods for initializing a land surface model, J. Hydrometeor., 6, 146-155.

Rushton K., 2007: Representation in regional models of saturated river–aquifer interaction for gaining/losing rivers. J. Hydrol., 334, 262–281.

Schaake, J. C., Q. Duan, V. Koren, K. E. Mitchell, P. R. Houser, E. F. Wood, A. Robock, D. P. Lettenmaier, D. Lohmann, B. Cosgrove, J. Sheffield, L. Luo, R. W. Higgins, R. T. Pinker, and J. D. Tarpley, 2004: An intercomparison of soil moisture fields in the

North American Land Data Assimilation System (NLDAS), J. Geophys. Res., 109, D01S90, doi:10.1029/2002JD003309.

Schär, C., D. Luthi, and U. Beyerle, 1999: The soil-precipitation feedback: A process study with a regional climate model. J. of Climate, 12, 722-741.

Schmidhalter, U., H. S. Salem, and J. J. Oertli, 1994: Measuring and modeling root water uptake based on 36 chloride discrimination in a silt loam soil affected by groundwater. Soil Science, 158, 97–105.

Schwarzkopf, M. D., and S. B. Fels, 1991: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. J. Geophys. Res., 96, 9075–9096.

Scott, R. L., T. E. Huxman, D. G. Williams, and D. C. Goodrich, 2006: Ecohydrological impacts of woody-plant encroachment: Seasonal patterns of water and carbon dioxide exchange within a semiarid riparian environment. Glob. Change Biol., 12, 311–324.

Serpetzoglou E., E. N. Anagnostou, A. Papadopoulos, E. I. Nikolopoulos, and V. Maggioni, 2010: Error Propagation of Remote Sensing Rainfall Estimates in Soil Moisture Prediction from a Land Surface Model. J. Hydrometeor, 11, 705–720, doi: 10.1175/2009JHM1166.1.

Serpetzoglou, E., T. S. Karacostas, I. Pytharoulis, P. Zanis, A. Papadopoulos, and E. N. Anagnostou: Sensitivity analysis on the ingestion of remotely sensed precipitation data into the land surface scheme of a mesoscale model. *Nat. Hazards Earth Syst. Sci.*, to be submitted.

Seuffert, G., P. Gross, C. Simmer, and E. F. Wood, 2002: The influence of hydrologic modeling on the predicted local weather: two-way coupling of a mesoscale weather prediction model and a land surface hydrologic model. J. Hydrometeor., 3, 505–523.

Shafer, M. A., C.A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma Mesonet. J. Atm. Ocean. Tech., 17, 474-494.

Skofronick-Jackson, G., W. A. Petersen, W. Berg, C. Kidd, E. F. Stocker, D. B. Kirschbaum, R. Kakar, S. A. Braun, G. J. Huffman, T. Iguchi, P. E. Kirstetter, C. Kummerow, R. Meneghini, R. Oki, W. S. Olson, Y. N. Takayabu, K. Furukawa, and T. Wilheit, 2017: The Global Precipitation Measurement (GPM) Mission for Science and Society. Bull. Amer. Meteor. Soc., doi:10.1175/BAMS-D-15-00306.1.

Smith, E. A., G. Asrar, Y. Furuhama, A. Ginati, A. Mugnai, K. Nakamura, R. F. Adler, M.-D. Chou, M. Desbois, J. F. Durning, J. K. Entin, F. Einaudi, R. R. Ferraro, R. Guzzi, P. R. Houser, P. H. Hwang, T. Iguchi, P. Joe, R. Kakar, J. A. Kaye, M. Kojima, C. Kummerow, K.-S. Kuo, D. P. Lettenmaier, V. Levizzani, N. Lu, A. V. Mehta, C. Morales, P. Morel, T. Nakazawa, S. P. Neeck, K. Okamoto, R. Oki, G. Raju, J. M. Shepherd, J. Simpson, B.-J. Sohn, E. F. Stocker, W.-K. Tao, J. Testud, G. J. Tripoli, E. F. Wood, S. Yang, and W. Zhang, 2007: International Global Precipitation Measurement (GPM) program and mission: An overview. Measuring Precipitation from Space: EURAINSAT and the Future, V. Levizzani, P. Bauer, and F. J. Turk, Eds., Springer, 611–654.

Snyder, K. A., and D. G. Williams, 2000: Water sources used by riparian trees varies among stream types on the San Pedro River, Arizona. Agric. Forest Meteorol., 105, 227–240.

Sophocleous, M. A., 2002: Interactions between groundwater and surface water: the state of the science. Hydrogeology J., 10 (1), pp. 52–67.

Sorooshian, S., K.L. Hsu, X. Gao, H.V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN System Satellite–Based Estimates of Tropical Rainfall. Bull. Amer. Meteor. Soc., 81, 2035–2046.

Steiner A. L., J. S. Pal, S. A. Rauscher, J. L. Bell, N. S. Diffenbaugh, A. Boone, L. C. Sloan, and F. Giorgi., 2009: Land surface coupling in regional climate simulations of the West African monsoon. Climate Dynamics, 33(6), 869-892.

Steinwand, A. L., R. F. Harrington, and D. Or, 2006: Water balance for Great Basin phreatophytes derived from eddy covariance, soil water, and water table measurements. J. Hydrol., 329, 595–605.

Stieglitz, M., D. Rind, J. Famiglietti, and C. Rosenzweig, 1997: An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling. J. Climate, 10, 118–137.

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

Teuling, A. J., R. Uijlenhoet, F. Hupet, E. E. van Loon, and P. A. Troch, 2006: Estimating spatial mean root-zone soil moisture from point-scale observations. Hydrol. Earth Syst. Sci., 10, 755-767.

Tian, Y., C. D. Peters-Lidard, B. J. Choudhury, and M. Garcia, 2007: Multitemporal Analysis of TRMM-Based Satellite Precipitation Products for Land Data Assimilation Applications. J. Hydrometeor., 8, 1165–1183.

Tian, X., Z. Xie, and A. Dai, 2008: A land surface soil moisture data assimilation system based on the dual-UKF method and the Community Land Model, J. Geophys. Res., 113, D14127, doi:10.1029/2007JD009650.

Viterbo, P., 1995: Initial values of soil water and the quality of summer forecasts. ECMWF Newslett., 69, 2–8.

Walker, J.P., P. Houser, and R. Reichle, 2003: New technologies require advances in hydrologic data assimilation, EOS, 84, 545.

Walker, J. P., and P. Houser, 2004: Requirements of a global near-surface soil moisture satellite mission: accuracy, repeat time, and spatial resolution. Adv. Water Resour., 27, 785-801.

Walko, R. L., L. E. Band, J. Baron, T. G. F. Kittel, R. Lammers, T. J. Lee, D. Ojima, R. A. Pielke Sr., C. Taylor, C. Tague, C. J. Tremback, and P. L. Vidale, 2000: Coupled atmosphere-biophysics-hydrology models for environmental modeling. J. Appl. Meteorol., 39, 931–944.

Wan, Z. M., 2008: New refinements and validation of the MODIS land-surface temperature/emissivity products. Rem. Sens. Environ., 112, 59–74.

Wang D., G. Wang, and E. N. Anagnostou, 2005: Use of Satellite-Based Precipitation Observation in Improving the Parameterization of Canopy Hydrological Processes in Land Surface Models. J. Hydrometeor., 6, 745–763.

Wang D., E. N. Anagnostou, and G. Wang, 2006: The effect of sub-grid rainfall variability on the water balance and flux exchange processes resolved at climate scale: the European region contrasted to Central Africa and Amazon rainforests. Adv. Geosci., 7, 269–274.

Wang D., G. Wang, and E. N. Anagnostou, 2009a: Impact of sub-grid variability of precipitation and canopy water storage on hydrological processes in a coupled land–atmosphere model. Climate Dynamics, Vol. 32, 649-662, DOI: 10.1007/s00382-008-0435-1.

Wang, L., T. Koike, K. Yang, P. J.-F. Yeh, 2009b: Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin. J. Hydrol., doi:10.1016/j.jhydrol.2009.08.005.

Weaver C. P., 2006: Coupling between large-scale atmospheric processes and mesoscale land-atmosphere interactions in the US Southern Great Plains during summer. Part I: Case studies. J. Hydrometeor., 5(6), 1223-1246.

Wedgbrow, C. S., R. L. Wilby, H. R. Fox, and G. O'Hare, 2002: Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales. Int. J. Climatol., 22 (2), 219-236.

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences: An Introduction. Academic Press, 467 pp.

Xie, Z. H., and X. Yuan, 2010: Prediction of water table under stream-aquifer interactions over an arid region. Hydrol. Process., 24 (2), 160-169.

Yang, Z. L., and G. Y. Niu, 2003: The versatile integrator of surface and atmosphere processes—Part 1: model description, Global Plan. Change, 38(1–2), 175–189.

Yeh, P. J.-F., and E. A. B. Eltahir, 2005a: Representation of water table dynamics in a land surface scheme. Part I: Model development. J. Climate, 18, 1861–1880.

Yeh, P. J.-F., and E. A. B. Eltahir, 2005b: Representation of water table dynamics in a land surface scheme. Part II: Subgrid variability. J. Climate, 18, 1881–1901.

Yeh, P. J. F., and J. S. Famiglietti, 2009: Regional Groundwater Evapotranspiration in Illinois. J. Hydrol., 10(2), 464-478.

York, J. P., M. Person, W. J. Gutowski, and T. C. Winter, 2002: Putting aquifers into atmospheric simulation models: An example from the Mill Creek Watershed, northeastern Kansas. Adv. Water Resour., 25, 221–238.

Yuan X., Z. H. Xie, J. Zheng, X. J. Tian, and Z. L. Yang, 2008: Effects of water table dynamics on regional climate: A case study over east Asian monsoon area. J. Geophys. Res. Atmos., 113(d21), D21112.

Zampieri, M., F. D'Andrea, R. Vautard, Ph. Ciais, N. de Noblet-Ducoudré, and P. Yiou, 2009: Hot European Summers and the role of soil moisture in the propagation of Mediterranean drought. J. Climate, 22, 4747–4758.

Zampieri, M., E. Serpetzoglou, E. N. Anagnostou, E. I. Nikolopoulos, and A. Papadopoulos, 2012: Improving the representation of river–groundwater interactions in land surface modeling at the regional scale: Observational evidence and parameterization applied in the Community Land Model. J. Hydrol., 420–421, 72-86.

Zhang W. H., and D. R. Montgomery, 1994: Digital elevation model grid size, landscape representation, and hydrologic simulation. Water Resour. Res., 30(4), 1019-1028.

Zhang, H., and C. S. Frederiksen, 2003: Local and non-local impacts of soil moisture initialization on AGCM seasonal forecasts: A model sensitivity study. J. Climate, 16, 2117– 2137.

Zhao, Q., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. Mon. Wea. Rev., 125, 1931–1953.