Doctoral Thesis

Observational and modelling investigations of land-precipitation coupling

Author(s):
Guillod, Benoît P.

Publication Date:
2014

Permanent Link:
https://doi.org/10.3929/ethz-a-010211745

Rights / License:
In Copyright - Non-Commercial Use Permitted
Observational and modelling investigations of land-precipitation coupling

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

Benoît P. Guillod

MSc ETH Environ. Sc, ETH Zurich

born April 4, 1986

citizen of

Lausanne VD and Bas-Vuilly FR

accepted on the recommendation of

Prof. Dr. Sonia I. Seneviratne, examiner
Dr. Boris Orlowsky, co-examiner
Dr. Christopher M. Taylor, co-examiner

2014
Abstract

The land surface is an important component of the climate system. Via interactions with the atmosphere, it impacts temperature and precipitation. In particular, soil moisture influences the partitioning of the energy at the land surface into the sensible ($H$) and latent ($\lambda E$) heat fluxes, mainly in transitional regions between wet and dry climates. This partitioning can be expressed by the Evaporative Fraction (EF), the ratio of $\lambda E$ to the available energy. EF directly impacts temperature via $H$ (soil moisture-temperature coupling). The coupling between the land surface and precipitation works through several physical mechanisms. These include “direct” moisture recycling, that is via the input of water to the atmosphere through $\lambda E$, and “indirect” impacts of both $H$ and $\lambda E$ on the boundary layer growth and stability (local temporal coupling), as well as mesoscale circulations induced by spatial soil moisture gradients (spatial coupling).

The aims of this thesis are the following: First, I investigate the role of soil textures for land-atmosphere interactions over Europe in a regional climate model (RCM). Soil texture can control soil moisture, its relation to EF and thereby climate. Second, I statistically investigate the land-precipitation coupling in different observation-based datasets. This is first done over North America, where the local temporal coupling and its sources are explored, and then in global analyses of temporal and spatial soil moisture-precipitation couplings.

In the first part of this thesis, regional climate simulations over Europe with two different maps of soil texture are compared. Soil textures typically define a number of thermal and hydraulic soil properties in a land surface model. Impacts of the choice in soil map on the simulated mean summer climate are large in some regions: up to 2°C in 2-meter temperature and 20% in precipitation. This effect is comparable to differences between simulations from different RCMs and it highlights the importance of soil properties for climate simulations. Furthermore, I identify the most important soil parameters for these effects (field capacity, plant wilting point, hydraulic diffusivity) and show the importance of the vertical profile of soil moisture in these simulations.
For the second part, investigations of land-precipitation coupling over North America reveal large uncertainties from different datasets. Results from a previous study, which highlights a positive EF-precipitation coupling over the Eastern US based on a reanalysis product, cannot be confirmed from the other observation-based data products. Instead, from remote-sensing products I find a positive coupling over the Central and Western US. Analyzing the drivers of the coupling signal shows that precipitation persistence plays a major role and prevents a clear identification of a causal relationship. Moreover, it suggests that the reanalysis-based signal in the Eastern US is driven by unrealistic evaporation of vegetation interception as well as atmospheric controls on EF. In remote-sensing products, the positive signal in the Central and Western US is likely related to soil moisture. These findings solve some of the discrepancies between contradicting results from the literature.

At the global scale, I find widespread negative spatial coupling, consistent with previous studies. This is contrasted by a dominance of positive temporal relationships in most regions, where precipitation events occur more often when soils are wet, while being located over comparatively drier patches. If the temporal relationship depict an actual coupling, this suggests that spatial and temporal coupling metrics quantify different processes. One may expect the land surface to amplify and perpetuate temporal anomalies and, at the same time, to reduce spatial soil moisture gradients. Further work will investigate the processes underlying the respective coupling signals and the origin of the temporal signal.

The results of this thesis shed light on some of the open questions about the coupling between the land surface and precipitation. They offer insights into the apparent contradictions of several recent studies on the topic, opening paths of future research. They also underline the complexity of the involved physics and thereby make a case for the necessity of better and more comprehensive datasets. The ongoing improvement of remote sensing data products together with the herein proposed findings and interpretations could over the next years contribute to a deeper and better constrained understanding of the interactions between the land surface and precipitation.
Résumé

La surface continentale est une importante composante du système climatique, qui peut influencer la température et les précipitations par ses interactions avec l’atmosphère. En particulier, l’énergie disponible à la surface est retournée à l’atmosphère sous forme de flux de chaleur sensible ($H$) et latente ($\lambda E$). Le partitionnement entre $H$ et $\lambda E$, qui peut être quantifié par la fraction évaporative (EF), c’est-à-dire le rapport entre $\lambda E$ et l’énergie disponible, est influencé par l’humidité du sol, principalement dans les régions de transition entre climats secs et humides. EF influence la température directement via $H$ (le couplage entre l’humidité du sol et la température). Le couplage entre la surface et les précipitations est plus complexe et fonctionne à travers différents mécanismes. Ceux-ci incluent un effet direct nommé “recyclage de l’humidité”, via la contribution de $\lambda E$ à l’humidité de l’atmosphère, et des effets indirects impliquant $H$ et $\lambda E$: Les flux de surface peuvent influencer la formation des précipitations via leur effet sur la croissance de la couche limite et la stabilité de l’atmosphère (couplage temporel local) ainsi que via des circulations à méso-échelle créées par des gradients spatiaux d’humidité du sol (couplage spatial).

Les buts de cette thèse sont les suivants: Tout d’abord, de tester l’importance des texture du sol pour les interactions entre la surface et l’atmosphère en Europe dans un modèle régional climatique (RCM). La texture du sol peut contrôler l’humidité du sol, sa relation à EF et, ainsi, le climat. Ensuite, d’étudier à l’aide d’outils statistiques le couplage entre la surface et les précipitations, à partir de données de mesure. Une première analyse se concentre sur l’Amérique du Nord, où le couplage temporel local et ses contrôles sont examinés. Puis, des analyses et comparaisons des couplage temporels et spatiaux sont appliquées à l’échelle globale.

Dans la première partie de cette thèse, des simulations régionales climatiques utilisant différentes cartes de textures du sol sont comparées. En pratique, les textures du sol définissent un nombre de paramètres thermiques et hydrauliques du sol dans un modèle de surface. Les effets du choix d’une carte du sol sur le climat moyen simulé en été sont conséquents dans certaines régions: jusqu’à 2°C en température et 20% en précipitations. Ces effets sont
comparables aux différences entre des simulations de modèles climatiques différents, et soulignent l’importance des caractéristiques du sol pour les simulations du climat. De plus, j’identifie les paramètres du sol qui contribuent le plus aux effets observés (capacité au champ, point de flétrissement permanent, diffusivité hydraulique) et je démontre l’importance du profile vertical de l’humidité du sol dans ces simulations.

Dans la deuxième partie, des études du couplage entre la surface et les précipitations en Amérique du Nord révèlent de larges incertitudes liées aux jeux de données. Les résultats d’une étude précédente, qui a identifié un couplage positif à l’Est des États-Unis dans une réanalyse, ne peuvent pas être confirmé à l’aide d’autres données de mesure. Au contraire, sur la base de mesures par satellite, je trouve un couplage positif au Centre et à l’Ouest des États-Unis. L’analyse détaillée des facteurs contribuant au couplage montre que la persistance des précipitations joue un rôle majeur et empêche d’identifier clairement une relation de cause à effet. De plus, le signal trouvé sur l’Est des États-Unis est contrôlé par une surestimation de l’évaporation de l’eau interceptée par la végétation ainsi que par les contrôles de l’atmosphère sur EF. Dans les produits de mesures satellite, le signal positif trouvé au Centre et à l’Ouest des États-Unis est probablement lié à l’humidité du sol. Ces découvertes expliquent certains des désaccords entre différents résultats de la littérature.

À l’échelle globale, je trouve un couplage spatial négatif clairement dominant, comme souligné dans des études précédentes. Ce résultat contraste avec une domination de relations positives temporellement dans la plupart des régions, où les événements de pluie se produisent ainsi le plus souvent lorsque les sols sont humides mais sont localisés au-dessus de parcelles plus sèches. Si la relation temporelle correspond à une relation de cause à effet, ces résultats suggèrent que les métriques de couplage temporel et spatial quantifient différents processus. On peut s’attendre à ce que la surface amplifie et perpétue les anomalies temporelles, tout en réduisant simultanément les gradient spatiaux d’humidité du sol. La suite de ce travail va analyser les processus liés aux couplages respectifs et l’origine du signal temporel.

Les résultats de cette thèse éclaircissent certaines des questions ouvertes au sujet du couplage entre l’humidité du sol et les précipitations. Ils offrent des aperçus sur les contradictions apparentes entre différentes études sur le sujet, ouvrant ainsi de nouvelles perspectives de recherche. Ils soulignent également la complexité des processus physiques impliqués, ainsi que le besoin de jeux de données plus précis et plus complets. Dans les prochaines années, l’amélioration continue de produits de données par satellite ainsi que les découvertes et interprétations proposées ici pourraient contribuer à une compréhension plus affinée et plus précise du couplage entre la surface et les précipitations.
# Contents

Abstract ........................................... i

Résumé ........................................... iii

1 Introduction ...................................... 1
   1.1 The Water and Energy Balance ................. 2
   1.2 Observations and models ....................... 7
   1.3 Land-climate interactions ..................... 12
   1.4 Soil moisture-precipitation feedbacks ......... 18
   1.5 Aims and outline ................................ 27

2 Impact of soil map specifications for European climate simulations .......... 29
   2.1 Introduction .................................... 30
   2.2 Experimental setup .............................. 31
   2.3 Results ......................................... 38
   2.4 Discussion and conclusions .................... 50

3 Land-surface controls on afternoon precipitation diagnosed from observational data: uncertainties and confounding factors ................. 55
   3.1 Introduction .................................... 57
   3.2 Data sets ....................................... 60
   3.3 Methods ........................................ 69
   3.4 TFS from different data sets .................. 72
   3.5 Impact of EF vs precipitation persistence .... 78
   3.6 Soil moisture and interception evaporation .... 81
   3.7 Discussion and conclusions .................... 87
Climate change is increasingly recognized as a major threat to world’s peace and prosperity. Assessment reports from the Intergovernmental Panel on Climate Change (IPCC) have compiled the scientific basis of man-made climate change as well as its consequences (e.g. IPCC, 2007, 2013).

Of particular relevance to society, climate change is expected to impact the occurrence and strength of extreme events, such as heat waves, heavy precipitation events, and droughts (Seneviratne et al., 2012). The land surface can largely influence the magnitude of some extremes and changes thereof, in particular in the case of heat waves (Seneviratne et al., 2006b). While some of these processes are relatively well understood, the impacts of soil moisture on precipitation and possibly resulting feedbacks remain a scientific challenge (Seneviratne et al., 2010). This thesis presents a thorough investigation of interactions between the land surface and precipitation based on a variety of observations-based analyses and model experiments.

Since soils are intimately linked to the atmosphere through the land water and energy balances, this introduction starts with the water and energy cycles and balances, both at the global and at the land surface levels (Sec. 1.1). Data underlying our climate analyses (observations, models) are introduced in Sec. 1.2. I discuss land-climate interactions regarding soil moisture, soil properties and soil moisture-atmosphere feedbacks in Sec. 1.3, while soil moisture precipitation feedbacks are described in Sec. 1.4. The last section of this chapter (1.5) summarizes the aims and outline of this thesis.
1.1 The Water and Energy Balance

Water and energy are important quantities in the climate system (e.g. Wallace and Hobbs, 2006). This section presents a short overview of the global energy and water balance, with an emphasis on the land water and energy balance.

1.1.1 The Global Energy Budget

Figure 1.1 displays the global energy budget for the present climate. A large part of energy exchanges consists of radiation fluxes, both in the shortwave spectrum (or solar radiation, in yellow, left part of Fig. 1.1) and in the longwave spectrum (or thermal radiation, in orange, right part of Fig. 1.1). Shortwave radiation, provided to the Earth by the sun, is the external driver of climate. About 45% of the incoming shortwave radiation is either absorbed by molecules in the atmosphere or reflected by clouds and aerosols. Part of the remaining energy is reflected at the land surface and the rest is absorbed by the Earth’s surface. In contrast, longwave radiation is emitted by the Earth’s surface as a function of its temperature; indeed, the Earth
1.1. THE WATER AND ENERGY BALANCE

can be considered as a black body and thereby emits longwave radiation according to Stefan-Blotzmann law. Similarly, the atmosphere itself emits longwave radiation, part of which is going out of the Earth’s system to space. A substantial part of the upward longwave energy flux, however, is absorbed by greenhouse gases (water vapour, CO$_2$, methane, ozone and other gases), which subsequently re-emit it in all directions, including back to the surface. This greenhouse effect is crucial to life as it leads to a global mean temperature of about 15$^\circ$C and thus also allows for liquid water (Wallace and Hobbs, 2006). Nonetheless, a large amount of longwave energy in not reflected back to the Earth and leaves the Earth’s system.

In equilibrium, the net outgoing longwave radiation should be equal to the net incoming shortwave radiation. As shown in Fig. 1.1, this is not the case under present conditions. This imbalance in the top of atmosphere radiation budget, positive downwards, is a consequence of anthropogenic forcing, most importantly due to greenhouse gases emissions (IPCC, 2007). The surplus of energy is absorbed by the climate system and thereby leads to the currently observed global warming.

At the land surface, not all of the energy provided by the sun is emitted back as longwave radiation. The remaining amount of energy is transmitted to the atmosphere as evaporation and sensible heat. This is discussed in detail in Section 1.1.3, which addresses the land water and energy balances.

1.1.2 The Water Cycle

Figure 1.2 shows a schematic representation of the water cycle. Unlike energy, water cannot leave the Earth’s system; instead, it is constantly exchanged between a number of pools. The oceans contain most of the Earth’s water and provide a large amount of water to the atmosphere, which rains out either back to the oceans or to the land. On the land, water either runs off to rivers, lakes and oceans, infiltrates to groundwater, or is evaporated back to the atmosphere – from the soil, interception surfaces such as leaves, or through vegetation transpiration. Water is also present in solid form as snow or ice, which eventually melts or sublimates.

While the oceans contribute more to the global-scale atmospheric water than the land (see respective evaporation fluxes in Fig. 1.2), land evaporation constitutes a main source of moisture for precipitation over land. This highlights that the classical view of the oceans driving the moisture input for land precipitation is largely exaggerated, as recycling of water on the land via evaporation and subsequent precipitation is very important as well. Indeed, about 65% of the water falling as precipitation over land is re-evaporated, based on the numbers in Fig. 1.2; because some of this moisture is transported to ocean areas, the total contribution of land evapotranspiration to continental precipitation is, however, somewhat smaller, and likely at around 40% (van der Ent et al., 2010).
Figure 1.2: The hydrological cycle. Estimates of the main water reservoirs, given in plain font in $10^3$ km$^3$, and the flow of moisture through the system, given in slant font ($10^3$ km$^3$/yr). From Trenberth et al. (2007).
1.1. THE WATER AND ENERGY BALANCE

Figure 1.3: Schematic of the land water balance (left) and land energy balance (right) for a given surface soil layer. \( \frac{dS}{dt} \) refers to the change in water content within the layer (soil moisture, surface water, snow; depending on the depth of the layer, this may include ground water changes), while \( \frac{dH_s}{dt} \) refers to the change of energy within the same layer. \( \text{SW}_{\text{net}} \) refers to the net shortwave radiation (\( \text{SW}_{\text{in}} - \text{SW}_{\text{out}} \)) and \( \text{LW}_{\text{net}} \) refers to the net longwave radiation (\( \text{LW}_{\text{in}} - \text{LW}_{\text{out}} \)). Note that \( \text{H}_2\text{O} \) and \( \text{CO}_2 \) refer to atmospheric water vapour and atmospheric \( \text{CO}_2 \) and their role as greenhouse gases. For simplicity other greenhouse gases are not indicated on the figure. From Seneviratne et al. (2010).

1.1.3 The Land Water and Energy Balances

At the land surface, water is stored in multiple reservoirs: soil moisture, surface water, canopy water, groundwater, snow, and ice cover. Precipitation provides water at the land surface (Fig. 1.3), which is partitioned into evaporation (\( E \)), surface runoff (\( R_s \)), drainage or subsurface runoff (\( R_g \)), and replenishing water storage (\( S \)). Therefore, the land water balance can be written as

\[
P = E + R_s + R_g + \frac{dS}{dt},
\]

(1.1)

where \( \frac{dS}{dt} \) is the change in water storage. A number of characteristics of the land surface, such as vegetation or soil properties, can impact the water balance and in particular the partitioning of precipitation into the other components.

The surface energy balance being already described in Section 1.1.1, I here only emphasize a few key points. Net radiation \( R_{\text{net}} \), composed of its net shortwave (\( \text{SW}_{\text{net}} \)) and longwave (\( \text{LW}_{\text{net}} \)) components, provides energy at the land surface. \( \text{SW}_{\text{net}} \) is determined by the incoming solar radiation reaching the ground and by the albedo (the fraction of energy that is reflected at the surface), while \( \text{LW}_{\text{net}} \) depends on the surface temperature and the amount
CHAPTER 1. INTRODUCTION

reflected down by the atmosphere. $R_{\text{net}}$ can be written as

$$R_{\text{net}} = SW_{\text{in}} - SW_{\text{out}} + LW_{\text{in}} - LW_{\text{out}}.$$  \hspace{1cm} (1.2)

This energy ($R_{\text{net}}$) is then distributed into three fluxes: while a usually small amount penetrates into the soil (ground heat flux $G$), most of the energy is converted into turbulent fluxes of sensible ($H$) and latent ($\lambda E$) heat to the atmosphere. Considering an infinitesimally thin layer of soil, the heat storage ($H_s$, and associated change $\frac{dH_s}{dt}$) is negligible and the land energy balance reads

$$R_{\text{net}} = \lambda E + H + G.$$ \hspace{1cm} (1.3)

Here, the latent heat flux ($\lambda E$) is the latent energy of the evaporated water $E$ ($\lambda = 2.26 \cdot 10^6$ J/kg is the latent heat of water). From this, it is evident that the land water and energy balances (Equations 1.1 and 1.3, respectively) are tightly coupled through evaporation. When the water balance leads to low moisture storage (due to e.g. deficits in precipitation), $E$ can become strongly limited. In this case, if the energy supply ($R_{\text{net}}$) is large, $H$ will strongly increase as water cannot be evaporated to increase $\lambda E$, leading to strong warming of the lower troposphere and, therefore, to high temperatures. The relevance of moisture availability for temperature emerges from this coupling between these two balances at the land surface.

The relatively simple water balance of Equation 1.1, however, does not distinguish between individual components of land evaporation. These are presented with more details in Fig. 1.4, and they are of high relevance to the work presented in this thesis. Before reaching the ground, precipitation can be intercepted by vegetation or impermeable surfaces (e.g., tar areas in cities). The intercepted water evaporates back to the atmosphere ($E_I$) in a relatively short time (typically within hours, depending on the water vapour saturation of the atmosphere), while the water that is not intercepted reaches the ground and either infiltrates into the soil or flows at the surface as surface runoff ($R_s$). Infiltrated water evaporates before penetrating deep into the soil ($E_{\text{soil}}$), drains through the soil to eventually end up as subsurface runoff ($R_g$), or is taken up by roots to be transpired back to the atmosphere ($E_{\text{trans}}$). Finally, processes can differ in the case of snow. Here I simply note the additional contribution from snow, which can sublimate ($E_{\text{snow}}$). Thus, land evaporation $E$ from Equation 1.1 can be decomposed into it components as

$$E = E_I + E_{\text{soil}} + E_{\text{trans}} + E_{\text{snow}}.$$ \hspace{1cm} (1.4)

The distinction between these components is an important feature, as they are reacting at very different time scales. Interception is usually happening within a few hours following a precipitation event. Soil evaporation is slightly slower, being coupled to surface soil moisture, with a typical time scale of one to a few days. Plant transpiration often acts on a much longer time scales due
1.2 Observations and Models

Figure 1.4: Processes within the land water balance. Compared to Fig. 1.3(left), the distinction between the components of land evaporation (surface evaporation, plant transpiration, interception evaporation, snow sublimation) is made explicitly. Adapted from Bonan (2008).

to its dependence on deeper soil moisture, with typical time scales of weeks to months. Throughout this thesis, the term “land evaporation” is used to refer to $E$, the sum of the individual components as shown in Equation 1.4; similarly, “latent heat flux” refers to $\lambda E$.

1.2 Observational datasets and models to investigate land-precipitation coupling

Climate is a complex system which requires a good understanding of its various components and processes. These can be studied through a variety of models and observations.

Observations include measurements at monitoring (e.g. weather) stations, measurements from field campaigns, or ground- or satellite-based remote-sensing data. They provide a basis for statistical analyses of the climate system as well as for the validation of models. However, observations often have limited spatial and/or temporal coverage (e.g. station data and satellite products, respectively).

Modelling is a powerful and widespread tool in the climate community,
CHAPTER 1. INTRODUCTION

which investigates the physical processes of the climate system by means of computer simulations. In particular, General Circulation Models (GCMs) and Regional Climate Models (RCMs) form the basis of climate projections from the IPCC (IPCC, 2007, 2013). In addition, they can be used for sensitivity studies that allow to isolate e.g. interactions between individual components, and to establish causal relationships, the latter being often difficult to demonstrate with the sole use of observations.

To overcome the irregular and sometimes sparse availability of observational data in space and/or time, reanalyses combine numerical models with observations to objectively generate synthesized estimates of the state of the system.

In this section, I describes these various tools. Section 1.2.1 provides an overview of observations of the climate system. Section 1.2.2 introduces climate models, while Section 1.2.3 describes reanalysis products, in which models are constrained by observations.

1.2.1 Observations: from ground measurements to remote-sensing products

Measurements of climate-relevant quantities and products derived from these can be classified into different types and have evolved over the past decades. Figure 1.5 illustrates the various types of observations used for weather and climate operational and research products.

First, direct measurements at surface stations provide a basis for climate monitoring, often considered as a “ground truth”. Traditionally, these include measurements of temperature, humidity, precipitation, wind and other variables measured at weather stations and supervised by national weather agencies. These sometimes extend back in time to before the existence of national weather services, with the collection of measurement from private or administrative entities (e.g. universities) at that time. These stations are located on continents. Complementary but less relevant for this thesis, direct marine observations are available from ships as well as moored and drifting buoys.

These direct, station-based measurements have been complemented by other variables (e.g. radiation fluxes) as well as specific stations that measure particular land-relevant quantities. For instance, FLUXNET stations (Baldocchi et al., 2001; Baldocchi, 2008) provide recent measurements of exchanges of water, energy and carbon between the land surface and the atmosphere.

Upper-air direct observations are provided by radiosondes attached to free-rising balloons measuring pressure, wind velocity, temperature and humidity up to heights of about 30km. Several hundreds of stations launch radiosoundes twice a day, providing a relatively good coverage worldwide. Complementarily, observations are collected from aircrafts during flights.
In addition to these direct sources of observations, indirect estimates are provided by various remote-sensing products (ground-based, airborne, or satellite-based). These are of two main types: active and passive sensors. Active sensors emit a signal and measure its echo after reflection by the sounded volume or area, while passive sensors measure a signal directly emitted by the volume/area.

In between satellites and weather stations, weather radars are ground-based, active sensors that allow for high-resolution spatial and temporal mapping of precipitation. These scan the atmosphere in all directions and at different beam angles to estimate reflectivity. Since reflectivity depends mainly on the number and size of hydrometeors, it is then converted to rainfall rate. These are used widely by weather agencies in Europe and North America, such as the Next Generation Weather Radar (NEXRAD) network in the US (Fulton et al., 1998; Young et al., 1999).

Remote-sensing observations from space by sensors onboard various polar orbiting and geostationary satellites have become a major component of Earth observations in recent years (Yang et al., 2013). These provide data about atmospheric features such as clouds, but also soundings of the atmosphere (e.g. aerosols), precipitation, and land surface (e.g. skin temperature, soil moisture, land cover) and ocean variables (e.g., ocean salinity, temperature), among others. Record length varies from a few years to several decades.
for some variables, with often relatively high temporal resolution. One of the main advantages of satellite remote-sensing products is their spatial scale, which is often more comparable to models and more representative than point-scale station data. Some products are not directly measured but rather derived from satellite data. For instance, the Global Land surface Evaporation: the Amsterdam Methodology (GLEAM, see Miralles et al., 2011b) used in this thesis takes advantage of a number of satellite remote-sensing products (radiation, precipitation, soil moisture, skin temperature, vegetation optical depth, snow) to estimate land evaporation.

Numerous products have been derived from these different types of measurements, and often using combinations of these. For instance, gridded temperature and precipitation data at monthly and daily resolutions are widely established (e.g., for precipitation, CPC-Unified, Chen et al., 2008 and GPCP-1DD, Huffman et al., 2001) and their use may be more appropriate than the direct use of station data for applications such as model validation due to more comparable spatial scale.

1.2.2 Models

Climate models are useful tools for the investigation of the climate system and for climate projections. Basically, these consist of mathematical equations, discretized on a grid, which numerically describe the climate system. First-order physical principles, such as equations describing the motion and conservation of energy, mass and angular momentum as well as the thermodynamic equation, dictate the general circulation of the atmosphere. These are explicitly represented in a climate model.

GCMs and most RCMs grids are often rather coarse, with horizontal resolutions of about 100-250km and 10-50km, respectively. Therefore, they cannot explicitly account for smaller-scale processes, such as cumulus convection or turbulence. These small-scale processes are represented by parameterizations, based on physical understanding of the underlying processes or empirical knowledge.

The term GCM is generally used to designate an atmosphere-ocean coupled general circulation model (AOGCM), including a land surface model, or an Earth System Model (ESM), where the AOGCM is coupled to additional modules such as sea ice and atmospheric chemistry modules. Here I use “GCMs” in a general sense, including AOGCMs and ESMs.

Land surface models (LSMs) provide the lower boundary condition to the atmosphere and have become increasingly complex over the last decade. Their first generation solved the land water and energy balance (Sec. 1.1.3) in order to provide exchanges of energy (radiation, heat) and water with the atmosphere in a very simplistic way. Second-generation LSMS include biophysical principles where fluxes are coupled via stomatal conductance (Sellers et al., 1997). The current third-generation models represent complex
biogeochemical interactions within the plant and soil continuum related to photosynthesis and even dynamic vegetation. These models can also simulate the land surface in uncoupled mode, where they are driven by observations, reanalysis or model data.

GCMs are useful tools for global climate studies. However, their coarse resolution can be limiting, especially for studies focusing on a given region, as small-scale features such as complex topography or heterogeneous land surfaces and coastlines cannot be resolved properly. Therefore, RCMs have been used increasingly in the last decades (Giorgi, 2006). The higher resolution in RCMs compared to GCMs comes at the cost of not representing the whole globe but a limited area (e.g., Europe). Moreover, initial and lateral boundary conditions such as wind or pressure have to be provided, usually from GCMs or reanalyses.

**COSMO-CLM** In this thesis (Chapter 2), I use the non-hydrostatic RCM COSMO-CLM (Rockel et al., 2008). Developed in a cooperation between the COnsortium for Small-scale MOdeling (COSMO) and the Climate Limited-area Modeling Community (CLM-Community), it is based on the compressible non-hydrostatic equations of fluid dynamics, discretized on rotated geographical coordinates in the horizontal and terrain-following height coordinates in the vertical. Technical details are available in the online documentation (http://www.cosmo-model.org/content/model/documentation/core/default.htm). Here, I use version 4.8 of COSMO-CLM.

COSMO-CLM includes TERRA_ML (Schrodin and Heise, 2001; Grasselt et al., 2008), a second-generation LSM. This multilayer soil module describes heat transfer and water transport in the soil, as well as evaporation from bare soil and plant transpiration, where the latter is based on the empirical BATS model from Dickinson (1984), i.e. it is less physically-based than current third-generation LSMs and there is no explicit coupling with photosynthesis.

### 1.2.3 Reanalyses

Reanalysis products are produced by models constrained with observations via data assimilation (Trenberth et al., 2008). Such a system includes a forecast model (e.g. a GCM) and data assimilation routines to ingest observations (Bosilovich et al., 2008). Unlike routine analyses produced by operational meteorological centers in real time, reanalyses use a “frozen” version of the forecast model, avoiding spurious effects in time from the use of different model versions. Another advantage is that the amount of ingested observational data is greater, as many high-quality observations are not available in real time (e.g. Betts et al., 2006). These uniform, gridded datasets thereby contain temporally and spatially continuous data which are physically consistent with available observations.
In spite of these advantages, reanalysis products also suffer from several issues. While the model and the data assimilation scheme are frozen, observational datasets may change in time (Betts et al., 2006). Moreover, mass, energy and momentum are not conserved after the adjustment to observations. Finally, while variables directly constrained by observations (e.g. atmospheric humidity) are generally of good quality, other variables such as surface turbulent fluxes ($\lambda E$ and $H$) may contain substantial errors (Decker et al., 2011).

Out of the many existing reanalyses (e.g. Dee et al., 2013), I use two products in this thesis. The well-established ERA-Interim reanalysis (Dee et al., 2011) is a global reanalysis from the European Centre for Medium-range Weather Forecasts (ECMWF), covering the years 1979-present. The North American Regional Reanalysis (NARR, see Mesinger et al., 2006) is a recent product over North America. An advantage of NARR over other reanalyses is the assimilation of hourly precipitation data from station measurements with the aim of better constraining the land-surface, but large biases in other variables have been identified (e.g. surface radiation fluxes, see Kennedy et al., 2011), which may question the quality of the land surface representation in NARR.

1.3 Land-climate interactions

Climate and land surface processes interact in various ways. Exchanges of greenhouse gases, in particular CO$_2$ through photosynthesis and respiration from plants and other organisms, are an important component on long time scales and very relevant to climate change (IPCC, 2007). On shorter time-scales, however, exchanges of water and energy (Section 1.1) can be more relevant (e.g. Seneviratne et al., 2010). These are often linked to soil moisture via evaporation. In this section, I provide a short introduction about soil moisture and its properties (Section 1.3.1) followed by an overview of soil moisture-mediated interactions between land surface processes and the climate (Section 1.3.2).

1.3.1 Soil moisture definition and soil properties

There are a number of definitions of soil moisture, which most importantly differ in the depth or volume considered and the choice of units. Volumetric soil moisture ($\Theta$) is defined for a given volume of soil, $V_{soil}$, as

$$\Theta = \frac{\text{volume of water in } V_{soil}}{V_{soil}}.$$  \hspace{1cm} (1.5)

For a given sub-surface area parallel to the surface, Equation 1.5 can be rewritten using the height of water within that volume $h_{H_2O}$ and the height of
the soil volume $h_{\text{soil}}$ as $\Theta = h_{H_2O}/h_{\text{soil}}$. This is often used in models, where the soil is discretized in horizontal layers.

The choice of the height $h_{\text{soil}}$ (or volume $V_{\text{soil}}$) considered depends on the application. Usually, soil moisture is defined as the water contained in the unsaturated soil zone (Hillel, 1998; Seneviratne et al., 2010), while water in the saturated zone is termed “groundwater”. In this case, $h_{\text{soil}}$ would be the depth of the soil at which it becomes saturated with water. However, this depth can change and is not very convenient for many applications. Therefore, the depth actually considered might depend on the application. For instance, when analyzing soil evaporation, soil moisture in the top few cm of the soil may be most relevant, while in the presence of vegetation, $h_{\text{soil}}$ is often chosen to include the root-zone and thus better represent the water actually available to plants for transpiration. Note that surface soil moisture is part of the root-zone. In this thesis, both surface and root-zone soil moisture are considered. “Soil moisture” is used when referring to soil moisture in general and includes potentially the whole unsaturated zone, but mostly root-zone soil moisture. When referring to a particular depth of volume, I use “root-zone soil moisture” or “surface soil moisture”.

$\Theta$ is well defined for a given depth. Nonetheless, the units used ($m^3$ of water/$m^3$ of soil) may not always be suitable. Indeed, a number of soil and vegetation properties define the range over which $\Theta$ can vary. The porosity $\Theta_{PV}$ (volumetric fraction of air and water in the soil) defines the maximum potential value for $\Theta$. However, in reality, part of the water cannot be held by the soil against gravitational drainage; therefore, field capacity $\Theta_{FC}$, the amount of water remaining in the soil after removal of excess water by gravitational drainage, is more representative of the actual maximum amount of water that can be sustained for days. Similarly, in the context of water uptake by vegetation, the permanent wilting point $\Theta_{\text{PWP}}$ (or plant wilting point) is defined as the volumetric water content below which the water is held by the soil matrix too strongly to be accessible to plants (Hillel, 1998). Thus, one can express soil moisture relative to its natural range using a “soil moisture index” (SMI; see e.g. Betts, 2004, Seneviratne et al., 2010) defined as

$$\text{SMI} = \frac{\Theta - \Theta_{\text{PWP}}}{\Theta_{FC} - \Theta_{\text{PWP}}}. \quad (1.6)$$

SMI usually varies between 0 (no water available) and 1 (water not limiting).

Most of the various parameters defined in the previous paragraphs refer to local properties that can be highly variable in space: the root-zone is a property of vegetation, while $\Theta_{FC}$, $\Theta_{\text{PWP}}$ and $\Theta_{PV}$ depend on soil texture. Soil textures are usually classified into discrete classes based on the relative proportions of sand, clay and silt in the soil, using the soil texture triangle (Fig. 1.6). Typically in land-surface models, a number of soil parameters are defined as a function of soil textures classes, and one soil class is assigned to each grid point. The implicit assumption that a “mean” soil texture is
CHAPTER 1. INTRODUCTION

Figure 1.6: Soil texture triangle, which defines soil classes based on the relative amount of sand, clay and silt in the soil. Figure from http://www.soilsensor.com/images/soiltriangle_large.jpg (retrieved October 30, 2013).

representative for the whole grid cell (typically 20–50km in a regional climate model) may not be always valid; nonetheless, it uses information about soil texture to constrain parameters that have to be available at model scales.

Due to these highly variable parameters, expressing soil moisture relative to these properties, e.g. SMI computed over the root zone, is crucial when comparing soil moisture spatially and between different products (even when comparing different land-surface models, soil moisture is defined in various ways and objective comparison is not possible using $\Theta$ directly). Part of the uncertainties in soil moisture measurements and their spatial representativeness arises from such issues as well.

Measurements Soil moisture can be measured in various ways. In situ ground measurements are closest to the “ground truth” but are limited in spatial and temporal coverage (Seneviratne et al., 2010). They are not always representative of the larger-scale soil moisture, but they may be representative of large-scale soil moisture variability (Mittelbach and Seneviratne, 2012). In recent years, remote-sensing products based on active and passive sensors onboard various satellites (e.g. ASCAT and AMSR-E, see Wagner et al., 2013 and Owe et al., 2008, respectively) have provided global, large-scale (typically at a 0.25° resolution) and frequent (global coverage every 2 days) estimates of surface soil moisture (e.g. Loew et al., 2013). Although the quality of these products might not be as high as that of direct measurements and depends
on postprocessing algorithms, they provide global measurements at scales relevant to climate studies, albeit with short record lengths. Water-balance estimates such as the Basin-Scale Water Balance dataset (Seneviratne et al., 2004; Hirschi et al., 2006a,b; Mueller et al., 2011) provides estimates of total water storage based on the coupled land-atmosphere water balance at the scale of catchments based on reanalysis and runoff data. Finally, estimates from land-surface models are also available. For a in-depth review of these various methods, see Seneviratne et al. (2010).

**Soil moisture memory** Leaving aside the numerous issues related to soil moisture definition and measurements, soil moisture varies at time scales that are much slower than those of typical atmospheric variability, a property called “soil moisture memory” (i.e., persistence) (Koster and Suarez, 2001). Given the ability of soil moisture to impact the atmosphere via the water and energy balance, soil moisture persistence could provide persistence to the atmosphere (e.g. Koster and Suarez, 2001; Seneviratne and Koster, 2011; Orth and Seneviratne, 2012), similarly to sea surface temperature. Therefore, soil moisture is increasingly considered useful for forecasts initiation, as analyzed in-depth in the multi-model GLACE-2 experiment (Phase 2 of the Global Land-Atmosphere Coupling Experiment, see Koster et al., 2010, 2011).

### 1.3.2 Soil moisture in the climate system

As mentioned in Section 1.1.3, soil moisture impacts climate by controlling the $E$ component of the land water and energy balances. Here I discuss soil moisture-climate interactions, with emphasis on the coupling between soil moisture, evaporation and climate.

**Soil moisture-evaporation coupling** Soil moisture provides water available for evaporation (more specifically, soil evaporation and plant transpiration) and therefore impacts $E$ and $\lambda E$ directly. By modulating $\lambda E$, soil moisture also affects $H$ (Eq. 1.3) and thus both temperature and humidity in the atmosphere (Seneviratne et al., 2010). In other words, soil moisture can impact the partitioning of the energy available at the land surface ($R_{\text{net}} - G$) into $H$ and $\lambda E$, which can be quantified using the Evaporation Fraction (EF) as

$$ EF = \frac{\lambda E}{R_{\text{net}} - G} = \frac{\lambda E}{H + \lambda E}. $$ (1.7)

EF is preferred to the Bowen ratio $B_w = \frac{H}{\lambda E}$ (e.g. Crago and Brutsaert, 1996), because it is better defined for the limiting cases. $EF = 0$ ($B_w = \infty$) indicates that there is no evaporation and all the energy available is used by the sensible heat flux, while $EF = 1$ ($B_w = 0$) indicates that evaporation utilizes all the energy available, reducing $H$ to 0. EF is relatively constant during daytime...
EF = \frac{\lambda E}{R_{\text{net}} - G}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure17}
\caption{Definition of soil moisture regimes and corresponding EF regimes. EF denotes the evaporative fraction, and EF_{\text{max}} its maximal value. Adapted from Seneviratne et al. (2010).}
\end{figure}

(Crago and Brutsaert, 1996; Crago, 1996; Gentine et al., 2007, 2011a). For \( \lambda E = H \), EF = 0.5 while \( B_w = 1 \), hence \( B_w \) is not very convenient as it varies between 0 and 1 for \( H < \lambda E \) and between 1 and \( \infty \) for \( H > \lambda E \).

Evaporation is limited by both soil moisture and the atmospheric conditions constraining evaporative demand (in particular, the available energy, \( R_{\text{net}} - G \)) (e.g. Teuling et al., 2009a; Seneviratne et al., 2010). It is often expressed as

\[ E = S E_{\text{pot}}, \]

(1.8)

where \( E_{\text{pot}} \) denotes potential evaporation (i.e., evaporation when water supply is not limiting) and \( S \) is a factor representing soil moisture stress, ranging from 0 to 1. \( S \), sometimes termed \( \beta \)-factor (e.g. Sellers et al., 1997), is typically set to 1 when soil moisture is above a specific critical threshold (e.g. field capacity) and zero when it is below a critical threshold value (e.g. permanent wilting point), with a monotonous function for intermediate soil moisture values (e.g. linear interpolation between these values). Note that \( E_{\text{pot}} \) is to a large extent driven by \( R_{\text{net}} - G \). Thus, in a way EF normalizes \( E \) for the impact of radiation. Classically, conceptual frameworks based on Budyko’s work (Budyko, 1956, 1974) characterize the relationship between soil moisture and EF, as shown in Fig. 1.7. Three distinct regimes are defined: (i) a wet regimes, characterized by high soil moisture values (\( \Theta > \Theta_{\text{FC}} \)) and \( \text{EF} = \text{EF}_{\text{max}} \); (ii) a dry regime, characterized by low soil moisture values (\( \Theta < \Theta_{\text{PWP}} \) and no evaporation; and (iii) a transitional regime, characterized by a strong relationship between soil moisture and EF. A number
of studies investigate the geographical distribution of soil moisture-climate regimes (see e.g. Teuling et al., 2009a, and Seneviratne et al., 2010 for a review). Generally, transitional climate regions are located at the border of dry and wet regions. Evaporation is soil moisture limited in the dry and transitional regimes; it is energy-limited in the wet regime. $\text{EF}_{\text{max}}$ may be modulated by other processes, such as vegetation processes, and is not discussed in detail here.

In addition to the processes mentioned here, soil moisture also impacts the partitioning of rainfall into surface runoff and infiltration. Once the soil is saturated, water cannot infiltrate anymore and simply runs off. Below saturation, part of the water will infiltrate and thus replenish the soil moisture storage and, on longer time scales, groundwater storage.

**Soil moisture-temperature feedbacks** If soil moisture is limiting to EF, it directly impacts $H$ and, thereby, temperature. Soil moisture-temperature feedbacks have been shown to play a major role for summer temperature variability, both in present and future climate (e.g. Seneviratne et al., 2006b, 2010; Mueller and Seneviratne, 2012). Due to soil moisture memory, antecedent soil moisture conditions (e.g. in spring) can influence temperature and in particular the occurrence and strength of heat waves. This has been shown for the European climate (Vautard et al., 2007; Diffenbaugh et al., 2007; Hirschi et al., 2011) as well as globally (Mueller and Seneviratne, 2012). Quesada et al. (2012) demonstrate that, while certain atmospheric conditions are necessary for the development of heat waves, soil moisture deficits are an additional requirement to the occurrence of large heat waves in Central Europe. Humid soils tend to mitigate heat waves.

**Soil moisture-precipitation feedbacks** By modulating EF, soil moisture can also impact precipitation (Seneviratne et al., 2010). On the one hand, it controls the amount of evaporation $E$ and thereby the water available in the atmosphere for precipitation. On the other hand, it also impacts the energy available in the planetary boundary layer (BL), provided by turbulent fluxes of sensible and latent heat ($H$ and $\lambda E$). This latter effect impacts temperature and humidity in the BL and thus BL growth, which is crucial to the triggering of convection (e.g. Gentine et al., 2013). In addition, spatial gradients in EF may cause a mesoscale circulation (Taylor and Ellis, 2006) that could locally generate convection (Taylor et al., 2011).

Soil moisture-precipitation feedbacks are heavily debated in the literature (Seneviratne et al., 2010). A thorough overview of the current debate is provided in the next section (1.4).
1.4 Soil moisture-precipitation feedbacks

The scientific literature abounds in studies about soil moisture-precipitation feedbacks. These use different methods, data, concepts and tools, and often contradict one another for various reasons. Here, I discuss the basic terminology of the feedbacks, while next sections elaborate on the main physical mechanisms (Section 1.4.1). Literature reviews of three main physical mechanisms are presented in Sections 1.4.2 to 1.4.4.

Soil moisture-precipitation feedbacks can be represented as a closed loop between soil moisture, EF and precipitation. As shown in Fig. 1.8, the feedback loop is composed of three steps:

(A) Soil moisture-EF coupling: Soil moisture impacts EF (Section 1.3.2).

(B) EF-precipitation coupling: The moisture and heat input to the atmosphere corresponding to changes in EF impacts subsequent precipitation.

(C) Precipitation impacts soil moisture by replenishing soils.

Throughout this thesis, the term “feedback” refers to a closed feedback loop, while “coupling” refers to a one-way impact of a variable on another variable. In particular, “soil moisture-precipitation feedback” refers to the full loop shown in Fig. 1.8 (A-C), while “EF-precipitation coupling” refers to relationship B, i.e. the impacts of EF on precipitation.

Relationship A, the soil moisture-EF coupling, is most relevant in transitional regions between wet and dry climates, as discussed in Section 1.3.2. Here, I simply note the potentially negative feedback within relationship A since higher soil moisture content enabling higher EF (and thus higher $E$) leads to a faster depletion of soil moisture (e.g. Seneviratne et al., 2010). This tends to dampen an initial soil moisture anomaly. Therefore, for a positive soil moisture-precipitation feedback, the replenishing of soil moisture via increased precipitation (via A,B,C) must exceed the additional decrease in soil moisture resulting from the increased evaporation (e.g. Boé, 2013).

Relationship B, the EF-precipitation coupling, is complex as it involves a large number of processes. Even the sign of the coupling is subject to debate in the literature, with studies showing positive (e.g. Schär et al., 1999; Pal and Eltahir, 2001; Findell et al., 2011) or negative (Cook et al., 2006; Wei et al., 2008; Taylor et al., 2012) coupling, and even both coupling signs depending on e.g. atmospheric conditions (Findell and Eltahir, 2003a; Ek and Holtslag, 2004), regions (Findell and Eltahir, 2003b) or model parameterizations (Hohenegger et al., 2009). There are a number of positive and negative feedbacks within relationship B, which are not represented in the simplified framework of Fig. 1.8. For instance, EF may impact cloud cover, which impacts $R_{\text{net}}$ and therefore both $H$ and $\lambda E$, potentially providing sub-feedbacks within B.
Relationship C, i.e. the impact of precipitation on soil moisture, can be considered as direct and relatively well understood in most cases (Seneviratne et al., 2010). As discussed in Section 1.1.3, this depends on the partitioning of precipitation into interception storage, runoff, and infiltration into the soil.

Soil moisture-precipitation coupling (i.e., A-B) and in particular EF-precipitation coupling (B) still entail large uncertainties. These are the focus of the following sections.

### 1.4.1 Physical mechanisms

Soil moisture (or EF) can impact precipitation via a number of direct and indirect pathways. Here, I classify the various pathways that can be found in the literature into four categories of direct and indirect effects as follows:

(i) Direct effect: Moisture recycling

(ii) Indirect effect: Local Coupling

(iii) Indirect effect: Induced mesoscale circulation

(iv) Indirect effect: Impacts on large-scale circulation
Figure 1.9: Illustration of the three main pathways of soil moisture-precipitation coupling discussed. Figures from (a) Brubaker et al. (1993), (b) Ek and Holtslag (2004) and (c) Taylor et al. (2011).
Pathways (i-iii) are discussed in detail and are presented schematically in Fig. 1.9. Pathway (iv) is less relevant to this thesis and is only described briefly here.

Pathway (i), i.e. the direct pathway, is relatively simple: absolute moisture input in the atmosphere from $E$ provides additional water which falls back to the land as precipitation. “Moisture recycling” is the quantification of this process as the fraction of precipitation that comes from local $E$, for a given spatial scale.

The three other categories relate to indirect effects of soil moisture on precipitation. These often suggest that convective precipitation can be sensitive to the surface turbulent fluxes in various ways. Large-scale or stratiform precipitation (e.g. fronts), on the other hand, is expected to be largely determined by atmospheric circulation.

Local coupling (i.e., ii) has attracted strong attention in the past two decades (e.g. Schär et al., 1999; Pal and Eltahir, 2001; Findell and Eltahir, 2003a; Ek and Holtslag, 2004; Santanello et al., 2009; van den Hurk and van Meijgaard, 2010; Findell et al., 2011; and Seneviratne et al., 2010 for a review). It assumes indirect impacts of EF on boundary layer stability and growth and, thereby, precipitation formation. These effects are mostly local, as the critical quantity is not the moisture input provided by $E$ but its role (as well as the role of $H$) in modulating the boundary layer and, ultimately, convective processes, which occur locally. The water used for precipitation might be advected from far away (e.g. it might originate from ocean evaporation), but the local conditions determine when and where it is raining.

Induced mesoscale circulation (i.e., iii) has emerged in recent years as an important part of soil moisture-precipitation coupling (e.g. Taylor et al., 2011). In this case, spatial gradient in soil moisture (or EF) may generate surface and tropospheric pressure gradients that can drive a mesoscale circulation. This circulation in turn impacts atmospheric fluxes of moisture and can control the location and occurrence of precipitation events. Similar processes have been suggested with respect to land surface heterogeneity in general, in soil moisture or vegetation cover (e.g. Chen and Avissar, 1994; Pielke et al., 1998).

Finally, impacts via large-scale circulation (i.e., iv) could happen when strong gradients in soil moisture (or EF) impact atmospheric pressure and, thereby, circulation. For instance, Fischer et al. (2007a) suggested that dry conditions in summer 2003 over Europe may have strengthened the high pressure system prevailing during that event, which helped perpetuating the predominant circulation during the heat wave, and was not favorable to precipitation. Sensitivity modelling studies have also demonstrated an impact of soil moisture on large-scale circulation in extreme scenarios with removed land evaporation (Shukla and Mintz, 1982; Goessling and Reick, 2006).
22

CHAPTER 1. INTRODUCTION

Thus, this indirect effect might be important in some situation, such as large-scale extreme events (see Fischer et al., 2007a) or extreme sensitivity studies (Shukla and Mintz, 1982; Goessling and Reick, 2011). In this thesis, I focus on impacts at smaller scales and hence more emphasis is put on the three other categories of soil moisture-precipitation coupling (i-iii), while impacts via large-scale circulation are not described in more details here.

In practice, these different pathways often act simultaneously and they can be difficult to disentangle. Nonetheless, some studies have shown that moisture recycling and impacts on large-scale circulation may be less relevant than more local impacts (e.g. Wei and Dirmeyer, 2012). In particular, Guo et al. (2006) have shown that soil moisture-precipitation coupling is mainly acting on convective precipitation, which is mostly driven by local processes. Lee et al. (2012) suggest that transpiration not only contributes to precipitation in the Amazon but also reduces its variability, through combined indirect effects via local coupling and large-scale convergence. These pathways are described with greater detail separately in the next sections.

1.4.2 Direct effect: precipitation recycling

Historically, the first investigations of soil moisture-precipitation coupling focused on the concept of moisture recycling (Seneviratne et al., 2010). Figure 1.9(a) shows a schematic description of precipitation recycling: For a given domain, atmospheric moisture is provided by atmospheric incoming humid air and local \( E \). Part of it falls as precipitation while the rest flows out of the domain. Moisture recycling is usually defined as the ratio of precipitation that comes from local \( E \), i.e., \( \rho = P_m/P \) in Fig. 1.9(a) where \( P_m \) is the precipitation of locally evaporated water.

Early quantifications were based on simple assumptions (e.g., well-mixed conditions, see Brubaker et al., 1993; Savenije, 1995b,a; Eltahir and Bras, 1996) and led to estimates of typically 10-50%, depending on the region and the season considered.

Recently, approaches based on vertically-integrated water vapour balance (e.g. Bisselink and Dolman, 2008, 2009), Lagrangian methods (i.e. backward trajectories, see e.g. Sodemann and Zubler, 2010) and moisture-tagging (e.g. van der Ent et al., 2010; van der Ent and Savenije, 2011) have been applied to reanalyses or model products. These estimates are consistent with previous estimates, but they also highlight some issues related to the definition of moisture recycling. In a recent review, Gimeno et al. (2012) provide an overview of the sources of continental precipitation and the remaining challenges.

A main issue is the dependence of moisture recycling on the scale of the considered domain. Recent studies avoided this issue by simply computing the ratio of local precipitation that evaporated anywhere on land (van der Ent et al., 2010). Alternatively, the computation of metrics such as lengths and time scales of moisture recycling has been introduced (van der Ent and
1.4. SOIL MOISTURE-PRECIPITATION FEEDBACKS

Savenije, 2011), which quantify the distance and time over which moisture is recycled, respectively. A global average of 40% of land precipitation was found to come from land evaporation van der Ent et al. (2010, see also Section 1.1.2). Van der Ent and Savenije (2011) estimate length scales of recycling which vary between 500-2000km in the tropics, 3000-5000 in temperate climate zones, and over 7000km over deserts, with time scales of typically 3-20 days.

The role of forests has also been investigated in this context. For instance, Spracklen et al. (2012) demonstrated that, in the tropics, air that has passed over extensive vegetated areas produces at least twice as much rain than air that has passed over little vegetation. This suggests that future deforestation may greatly reduce precipitation in these regions via reduced moisture recycling.

1.4.3 Indirect effect: Local Coupling

A large number of studies have highlighted indirect soil moisture-precipitation (or EF-precipitation) coupling via the boundary layer growth and convection triggering. As depicted in Fig. 1.9(b), complex interactions between EF and the boundary layer impact primarily temperature and humidity in the boundary layer. This subsequently impacts stability, boundary layer growth and thereby entrainment, cloud formation, radiation and, ultimately, precipitation (Ek and Holtslag, 2004).

First studies which examined this set of processes and proposed theoretical frameworks for indirect coupling suggested two ways by which soil moisture could impact precipitation (Eltahir, 1998; Zheng and Eltahir, 1998; Pal and Eltahir, 2001). First, a radiative impact via albedo: Moist soils are darker, absorbing more radiation and thereby increasing $R_{\text{net}}$. Second, moister soils lead to higher EF and thereby cool and wetten the boundary layer (note that this second effect might ultimately lead to an additional increase in $R_{\text{net}}$ via longwave radiation). These two processes were then thought to lead to a positive soil moisture-precipitation coupling via increased energy in the boundary layer (e.g., via concepts such as the moist static energy, Pal and Eltahir, 2001). Some of these theoretical considerations were complemented by experiments with idealized coupled land surface-boundary layer models. Recent studies focus on the impact via EF, as impacts through albedo have been shown to be limited (Seneviratne et al., 2010).

Other modelling studies had identified negative coupling mechanisms, usually related to insufficient buoyancy over wet soils in the presence of capping inversions (Giorgi et al., 1996; Beljaars et al., 1996). While a few theoretical studies, which had already proposed frameworks where the impact of soil moisture on precipitation via EF may lead to negative coupling (De Ridder, 1997, 1999), had not received much attention, the work of Findell and Eltahir (2003a) demonstrated that the sign of the coupling may depend on
atmospheric conditions in the early morning, using radio-soundings and a boundary layer model. These conditions vary temporally and spatially, producing both positive and negative coupling in different regions (Findell and Eltahir, 2003b). This was further confirmed by Ek and Holtslag (2004), who showed that the sign of the coupling may depend on the atmospheric stability above the boundary layer. Simultaneously, the Global Land Atmosphere Coupling Experiment (GLACE) multi-model study of Koster et al. (2004) showed regions of strong coupling in models, highlighting the importance of the discussed mechanisms. An additional radiative feedback via cloud cover was found to be weaker than impacts of soil moisture via EF (Schlemmer et al., 2012).

Many studies have attempted to confirm these findings using observational data, but most of these have remained unsuccessful up to now (Seneviratne et al., 2010). Besides the limited availability of observational data, in particular for EF and soil moisture measurements at relevant scales, statistical analyses cannot establish causal relationships for various reasons.

First, the impacts of soil moisture (or EF) on precipitation (A-B or B in Fig. 1.8) are difficult to disentangle from the impact of precipitation on soil moisture (C) (Seneviratne et al., 2010). Indeed, precipitation persistence is a major confounder and several studies (Findell and Eltahir, 1997; D’Odorico and Porporato, 2004; D’Andrea et al., 2006) have been shown to suffer from this deficiency (Salvucci et al., 2002; Kochendorfer and Ramirez, 2005; Teuling et al., 2005; Daly et al., 2009; Wei et al., 2008). Note, however, that a positive feedback would itself induce precipitation persistence – this feature is precisely what makes these feedbacks attractive from a weather and seasonal prediction perspective (e.g. Koster et al., 2011). Therefore, explicitly accounting for precipitation persistence and distinguishing between atmospheric and land-induced precipitation persistence is difficult (Alfieri et al., 2008).

Second, a detected covariability between soil moisture (or EF) and precipitation may be due to a third influencing variable (e.g. sea surface temperature) which controls both soil moisture and precipitation (Notaro, 2008; Orlowsky and Seneviratne, 2010; Sun and Wang, 2012). Hence, a correlation between soil moisture and subsequent precipitation is a necessary but not a sufficient condition for establishing the existence of a coupling.

While modelling studies can avoid this problem and better isolate the effect of soil moisture (or EF) on precipitation, they also suffer from deficiencies. Studies show that soil moisture-precipitation feedbacks are exaggerated in some models (Koster et al., 2003; Ferguson et al., 2012). The GLACE experiment (Koster et al., 2004) demonstrates that soil moisture significantly impacts precipitation over many transitional climate regions in global climate models, but also reveals a large spread between results from different GCMs (Koster et al., 2006). This discrepancy is related both to step A (Guo et al.,
2006; Comer and Best, 2012) and B (Wei and Dirmeyer, 2010). Dirmeyer et al. (2006) highlight large biases in the GLACE simulations, not only in surface variables but also in local covariability of key atmospheric and land surface variables. The second phase of GLACE, which concentrates on the impacts of soil moisture initialization on subseasonal forecast skills and relies on more advanced models, exhibits rather weak soil moisture-precipitation coupling (Koster et al., 2010, 2011). Therefore, while modelling studies and in particular GLACE1 made a strong case for soil moisture-precipitation coupling, the wide range of coupling exhibited by different models and their often poor performance with respect to the most relevant processes prevent clear conclusions on the existence and the sign of the coupling (Seneviratne et al., 2010).

The considerable disagreement between models stems from the uncertainty of certain process representations. In particular, the parameterization of convection is not straightforward and often leads to large biases (e.g. Chaboureau et al., 2004; Guichard et al., 2004). The choice of a convective scheme can even determine the sign of the coupling (Hohenegger et al., 2009). Boundary layer schemes and land surface models have also been shown to impact the measured coupling (Santanello et al., 2009, 2011, 2012). Dirmeyer (2011) suggests that hot spot regions of soil moisture-evaporation coupling are compromised in reanalysis data due to the non-stationarity of the observing system, while sets of land surface models and GCMs show more consistent results.

Therefore, modelling studies identify positive coupling (Schär et al., 1999; Betts, 2004, 2009; Meng and Quiring, 2010a; Hauck et al., 2011; Mei et al., 2013; Berg et al., 2013), negative coupling (Giorgi et al., 1996; Beljaars et al., 1996; Cook et al., 2006; Williams et al., 2012), or coupling of both signs depending on atmospheric conditions (e.g. van den Hurk and van Meijgaard, 2010) or years (Asharaf et al., 2012).

Koster et al. (2003) provides first observational evidence for a positive coupling over the US Southern Great Plains. Later work, however, either confirms (Mei and Wang, 2012), contradicts (Ruiz-Barradas and Nigam, 2012) or moderates these findings (Meng and Quiring, 2010b). Other investigations in the US suggest negative coupling (Myoung and Nielsen-Gammon, 2010a), strong atmospheric controls (Myoung and Nielsen-Gammon, 2010b; Zhang and Klein, 2010), and atmospheric controls on the sign (Roundy et al., 2013) or strength (Aires et al., 2013) of the coupling.

At the global scale, an integrated analysis of the ERA-40 reanalysis, which accounts for seasonal cycle and precipitation persistence, highlights widespread positive coupling (Lam et al., 2007). Several other studies also suggest coupling in many regions (e.g. Zeng et al., 2010), some of which highlighting that predominant atmospheric stability may explain different signs of coupling in different regions (Ferguson and Wood, 2011; Westra et al., 2012).
Similarly, Boé (2013) distinguishes between positive and negative coupling situation using a weather regime approach.

The importance of free tropospheric stability via entrainment as well as capping inversion is repeatedly invoked, often based on coupled land surface-boundary layer models (van Heerwaarden et al., 2009; Santanello et al., 2009; Gentine et al., 2013) or large-eddy simulations (Huang and Margulis, 2011), suggesting that the sign of the coupling may depend on the atmospheric conditions.

Finally, Dirmeyer and colleagues study soil moisture-precipitation coupling under present and future climates, using the ECMWF operational model (Dirmeyer et al., 2012) and models from the Coupled Model Intercomparison Project 5 (CMIP5) (Dirmeyer et al., 2013). They find global increases in coupling strength, suggesting an increased relevance of soil moisture-precipitation feedback for future climate, and, therefore, the potential for improvements in climate projections linked to a better understanding of soil moisture-precipitation feedbacks. Note, however, that their approach suffers from the same limitations related to causality as observation-based analyses.

1.4.4 Indirect effect: Induced mesoscale-circulation

Mesoscale circulation induced by soil moisture was first diagnosed from satellite data over the Sahel region (Taylor and Ellis, 2006). The sign of this spatial coupling was suggested to be negative: deep convection was shown to form preferentially over drier patches close to wet soils, as a result of wind induced by pressure gradients due to differences in $H$. Further analyses, based on measurement campaigns, support this hypothesis (Taylor et al., 2007).

Modelling studies, using large-eddy simulations (van Heerwaarden and de Arellano, 2008; Brunsell et al., 2011) and regional climate models (Wolters et al., 2010) demonstrate spatial impacts of land surface gradients. Taylor et al. (2011) demonstrates these effects over the Sahel in the context of soil moisture precipitation coupling, and a global analysis highlights the strong dominance, globally, of negative spatial coupling (Taylor et al., 2012).

Nonetheless, and as highlighted by Koster (2011), there are large differences between the observed negative spatial coupling and possible temporal coupling. Some confusion has emerged recently, with some scientists interpreting the results from Taylor et al. (2012) as a negative coupling as such, without distinguishing between the notions of “spatial” and “temporal” coupling. While it is true that the results from Taylor et al. (2012) do not exclude negative temporal feedbacks, the analyzed quantities directly relate to spatial soil moisture gradients and not to their temporal structure. Co-existence of such spatial soil moisture-precipitation coupling with positive temporal coupling is not impossible in theory, although no evidence of such “co-habitation” of spatial and temporal feedbacks of differing signs has been
1.5 AIMS AND OUTLINE

As discussed in the previous subsections, land-atmosphere interactions play an important role in the climate system, via memory effects and impacts on extreme events such as droughts and heat waves. Not all of the involved processes are well understood. In particular, the feedback between soil moisture, EF and precipitation is one of the most uncertain component of land-atmosphere interactions (Seneviratne et al., 2010). To investigate these interactions, the use of both models and observations is advantageous. Therefore, for this thesis I first study and quantify the role of soil parameters in the representation of the European climate and land-atmosphere interactions in a RCM. Then, I focus on the soil moisture-EF-precipitation coupling. Using observational and reanalysis data, I investigate (i) the relevant processes and the role of land surface water storages (soil moisture, interception reservoir) in the coupling over North America and (ii) the differences between temporal and spatial coupling at the global scale.

The overarching aim of this thesis is to contribute to the understanding of land-atmosphere coupling, with a focus on soil parameters and soil moisture-precipitation coupling. This can be summarized into the following research questions:

- What is the sensitivity of a RCM to soil parameters, and what are the associated uncertainties?
- Do observations from ground measurements and satellite remote-sensing confirm NARR-based EF-precipitation coupling findings from a previous study?
- What is the role of the different land evaporation components in this coupling?
- Do temporal and spatial land-precipitation coupling investigate similar or different processes, and how do they relate to each other?

This thesis consists of five chapters and three appendices. Two published articles (Chapter 2 and 3, the latter published as a discussion paper) and an article in preparation (Chapter 4) form its core - all of which are included as self-contained scientific contributions. The three appendices refer to the three chapters, respectively. These chapters are shortly summarized as follows:

- Chapter 2: Impact of soil map specifications for European climate simulations (Guillod et al., 2013). This paper investigates the mean climate simulated by COSMO-CLM with soil texture maps from two different sources. It highlights large differences and suggests
CHAPTER 1. INTRODUCTION

that soil parameters may play a greater role in climate models than previously assumed. In addition, the role of individual soil parameters is investigated and important parameters are identified. The vertical profile of soil moisture is also shown to play an important role.

• Chapter 3: **Land-surface controls on afternoon precipitation diagnosed from observational data: uncertainties and confounding factors, (Guillod et al., 2014).** NARR-based results from a previous study (Findell et al., 2011), which identified a region of EF-precipitation coupling over the Eastern US, are compared to estimates from observation-based datasets (FLUXNET, GLEAM, NEXRAD). Results differ widely between datasets, but investigations of potential confounding effects from precipitation persistence and of the role of the individual components of $E$ (see Sec. 1.1.3) reveal interesting mechanisms. In NARR, the signal is dominated by atmospheric controls on EF and by unrealistic vegetation interception. In GLEAM, no significant signal is found over the Eastern US without introducing unrealistically high evaporation from vegetation interception, while a signal over the Western US appears to be related to soil moisture.

• Chapter 4: **Global analysis of soil moisture-precipitation coupling using remote-sensing data, (Guillod et al., in preparation).** In this chapter, I investigate the soil moisture-precipitation coupling at the global scale, using satellite remote-sensing data. Analysis of local temporal coupling versus spatial coupling (induced mesoscale circulation) suggests differing mechanisms. More specifically, precipitation events are found to occur in generally wet conditions (apparent positive temporal coupling), while being located over drier (or less wet) patches (negative spatial couplings).

• Chapter 5: **Conclusions and outlook.** I draw overall conclusions of this thesis and suggest possible future research topics in this final chapter of the thesis.
Impact of soil map specifications for European climate simulations

Clim. Dyn., 40, 123–141, doi:10.1007/s00382-012-1395-z *
Benoit P. Guillod¹, Edouard L. Davin¹, Chistine Kündig¹†, Gerhard Smitatek², and Sonia I. Seneviratne¹

Abstract Soil physical characteristics can influence terrestrial hydrology and the energy balance and may thus affect land-atmosphere exchanges. However, only few studies have investigated the importance of soil textures for climate. In this study, we examine the impact of soil texture specification in a regional climate model (RCM).

We perform climate simulations over Europe using soil maps derived from two different sources: the soil map of the world from the Food and Agricultural Organization (FAO) and the European Soil Database from the European Commission Joint Research Center (JRC). These simulations highlight

*This publication was slightly changed from its original version by adapting abbreviations to ensure consistency throughout this thesis.
†Current affiliation: EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland.
¹Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland.
²Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany.
the importance of the specified soil texture in summer, with differences of up to 2°C in mean 2-meter temperature and 20% in precipitation resulting from changes in the partitioning of energy at the land surface into sensible and latent heat flux.

Furthermore, we perform additional simulations where individual soil parameters are perturbed in order to understand their role for summer climate. These simulations highlight the importance of the vertical profile of soil moisture for evapotranspiration. Parameters affecting the latter are hydraulic diffusivity parameters, field capacity and plant wilting point.

Our study highlights the importance of soil properties for climate simulations. Given the uncertainty associated with the geographical distribution of soil texture and the resulting differences between maps from different sources, efforts to improve existing databases are needed. In addition, climate models would benefit from tackling unresolved issues in land-surface modeling related to the high spatial variability in soil parameters, both horizontally and vertically, and to limitations of the concept of soil textural class.

2.1 Introduction

Global and regional climate simulations are subject to large uncertainties. These uncertainties relate to, on the one hand, model formulation and, on the other hand, to input parameters used in models, in particular those describing surface characteristics. These parameters are generally linked to vegetation or soil characteristics. Many studies have focused on the role of vegetation properties (e.g. LAI, stomatal conductance, root depth) due to increasing interest in the impact of land cover change on climate (e.g. Bonan, 2008; Pitman et al., 2009), which led to more detailed and accurate maps of these properties (e.g. Lawrence and Chase, 2007). On the contrary, few studies have investigated the role of soil parameters (such as porosity, heat capacity or hydraulic conductivity) on climate, although it remains unclear which of plant or soil parameters can impact climate more strongly. While some studies have suggested that vegetation parameters are indeed important (Mölders, 2005), others have shown that soil parameters could matter as much (Osborne et al., 2004) or even more (Richter et al., 2004). Moreover, studies have shown that some soil physical properties, in particular infiltration rate, porosity or hydraulic conductivity, can change depending on e.g. crops, crop management (Uhland, 1950) and land clearing and use (Ghuman et al., 1991; Alegre and Cassel, 1996; Zimmermann et al., 2006). A few studies have already highlighted their non-negligible role for exchanges of water and energy at the surface (Anders and Rockel, 2009; Seneviratne et al., 2006a). In particular, soil properties influence soil moisture (SM) which plays a crucial role for summer climate in mid-latitude regions, through its memory and feedbacks to the atmosphere (Seneviratne et al., 2010). Most notably,
SM controls evapotranspiration (E) in these regions and, therefore, the partitioning of energy at the surface between sensible (H) and latent (λE) heat fluxes, directly influencing near-surface temperature and humidity. In spite of the recognized importance of soil moisture, several issues related to soil physical parameters remain unresolved.

First of all, soil parameters are usually assigned as attributes of soil classes that refer to the soil texture (Teuling et al., 2009b). However, the range of a soil parameter can be rather large, even within a given soil class; in fact, its variability within a soil class is often larger than its variability between the classes (Mölders, 2005; Teuling et al., 2009b). Moreover, most parameters are model dependent (Kahan et al., 2006). For instance, even a basic parameter such as the water-holding capacity is highly variable between state-of-the-art AGCMs despite being long recognized as a key parameter for land-climate interactions (Seneviratne et al., 2006a).

In addition to these issues linked to the attribution of parameter values to the different soil classes, several data bases of the geographical distribution of soil classes exist, of which the FAO soil map of the world (FAO/UNESCO, 1974) is the most commonly used global dataset (Smiatek et al., 2008). Nevertheless, alternative products exist at regional scale, some of which have a higher resolution and, in some cases, are based on more recent and detailed information; in Europe, such a product is the European Soil Database (ESDB), released by the European Commission Joint Research Center (JRC) (European Commission and the European Soil Bureau Network, 2004).

In the present study, we provide a detailed investigation of the role of soil parameters in regional climate simulations for the European continent. First, we compare simulations with a Regional Climate Model (RCM) with soil maps derived from two different sources to assess the potential impact of the soil map itself. We then analyze the impact of individual soil parameters on the local climate with the help of additional simulations to identify key parameters.

This paper is structured as follows: Section 2.2 describes the model (2.2.1) as well as the conducted experiments (2.2.2). Section 2.3 presents the results from the comparison of the impact of different soil maps on summer climate (2.3.1) and the role of individual soil parameters (2.3.2). Finally, the main findings are summarized and discussed in Section 2.4.

2.2 Experimental setup

2.2.1 COSMO-CLM

COSMO-CLM (Rockel et al., 2008) is a non-hydrostatic RCM developed jointly by the COnsortium for Small-scale MOdeling (COSMO) and the Climate Limited-area Modeling Community (CLM-Community). It is based on the compressible non-hydrostatic governing equations of fluid dynamics,
which are discretized on rotated geographical coordinates in the horizontal dimensions and terrain-following height coordinates in the vertical. A detailed technical documentation is available at http://www.cosmo-model.org/content/model/documentation/core/default.htm. In this study, we use version 4.8 of COSMO-CLM. This model version has been used by Davin and Seneviratne (2012), who have shown that it includes a significant reduction of model biases compared to COSMO-CLM version 4.0.

The LSM used in this version of the model is TERRA_ML, a multilayer soil model parameterizing evapotranspiration of plants and bare soil as well as heat transfer and water transport in the soil (see Appendix A.1 and Schrodin and Heise, 2001; Grasselt et al., 2008). TERRA_ML is a second-generation land surface model, where transpiration is modeled without explicit coupling with photosynthesis. More specifically, the evapotranspiration parameterization is derived from the BATS model (Dickinson, 1984), which is more empirical and less physically based than current third-generation LSMs.

Most relevant for this study is the dependence of the parameterization in TERRA_ML on a set of parameters defined for eight different soil classes representing a wide range of soil textures, including two special classes (ice, rock) and six standard classes (sand, sandy loam, loam, loamy clay, clay and peat). 15 parameters are associated to each soil class and their respective values are shown in Table 2.1.

2.2.2 Experiments

In this study, we analyze a set of simulations with COSMO-CLM. These simulations differ only in the applied soil maps and/or in the look-up table of the soil parameters corresponding to each soil class. The rest of the setup is the same throughout the experiments, including the atmospheric part of the model. This allows to strictly isolate the effect of soil maps and/or soil parameters.

The configuration applied to all simulations is the following: The model is run over the European continent, including parts of North Africa and Western Russia (e.g. Fig. 2.1), with a horizontal resolution of $0.44^\circ$ ($\sim 50$ km), 32 vertical levels and a time step of 240s. Initial and lateral boundary conditions are based on ERA-40 reanalysis data, except for the years 2002-2005 where ECMWF operational forecast analyses are used. The model is run for the period 1980-2005, where the first six years serve as a spin-up in order to reach equilibrium, in particular to allow soil moisture to adjust to the modified conditions. The remaining 20 years (1986-2005) are used in the analysis. Although no ensemble runs were performed, the internal variability of the model is relatively well sampled given the simulated length.
### Table 2.1: Look-up table of soil parameters for each soil class in TERRA_ML. Values of the available water capacity $\theta_A = (\theta_{\text{FC}} - \theta_{\text{PWP}})$ are given as well. From Doms et al. (2011).

<table>
<thead>
<tr>
<th>Soil Type</th>
<th>Volume of voids $\theta_{PV}$ [-]</th>
<th>Field capacity $\theta_{\text{FC}}$ [-]</th>
<th>Permanent wilting point $\theta_{\text{PWP}}$ [-]</th>
<th>Air dryness point $\theta_{\text{ADP}}$ [-]</th>
<th>Minimum infiltration rate $I_{k2}$ [kg/(m$^2$ s)]</th>
<th>Hydraulic diffusivity $D_0$ [$10^{-9}$ m$^2$/s]</th>
<th>Hydraulic conductivity $K_0$ [$10^{-9}$ m/s]</th>
<th>Heat capacity $\rho c_0$ [$10^6$ J/(m$^3$ K)]</th>
<th>Heat conductivity $\lambda_0$ [W/(Km)]</th>
<th>Exponent $B$ [-]</th>
<th>Albedo [-]</th>
<th>Wet albedo [-]</th>
<th>Available Water Capacity $\theta_A = (\theta_{\text{FC}} - \theta_{\text{PWP}})$ [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ice</td>
<td>-</td>
<td>-</td>
<td>0.364</td>
<td>0.196</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.920</td>
<td>2.620</td>
<td>1</td>
<td>0.7</td>
<td>0.44</td>
<td>0.154</td>
</tr>
<tr>
<td>rock</td>
<td>-</td>
<td>-</td>
<td>0.445</td>
<td>0.260</td>
<td>0.042</td>
<td>0.012</td>
<td>-</td>
<td>2.100</td>
<td>0.300</td>
<td>3.5</td>
<td>0.3</td>
<td>0.27</td>
<td>0.160</td>
</tr>
<tr>
<td>sand</td>
<td>-</td>
<td>-</td>
<td>0.455</td>
<td>0.340</td>
<td>0.100</td>
<td>0.030</td>
<td>-</td>
<td>1.280</td>
<td>0.280</td>
<td>4.8</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td>SL</td>
<td>-</td>
<td>-</td>
<td>0.475</td>
<td>0.370</td>
<td>0.110</td>
<td>0.035</td>
<td>-</td>
<td>1.350</td>
<td>0.250</td>
<td>6.1</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td>L</td>
<td>-</td>
<td>-</td>
<td>0.475</td>
<td>0.370</td>
<td>0.110</td>
<td>0.035</td>
<td>-</td>
<td>1.420</td>
<td>0.250</td>
<td>6.1</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td>LC</td>
<td>-</td>
<td>-</td>
<td>0.475</td>
<td>0.370</td>
<td>0.110</td>
<td>0.035</td>
<td>-</td>
<td>1.500</td>
<td>0.250</td>
<td>6.1</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td>loamy clay</td>
<td>-</td>
<td>-</td>
<td>0.475</td>
<td>0.370</td>
<td>0.110</td>
<td>0.035</td>
<td>-</td>
<td>1.500</td>
<td>0.250</td>
<td>6.1</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td>clay</td>
<td>-</td>
<td>-</td>
<td>0.475</td>
<td>0.370</td>
<td>0.110</td>
<td>0.035</td>
<td>-</td>
<td>1.500</td>
<td>0.250</td>
<td>6.1</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td>peat</td>
<td>-</td>
<td>-</td>
<td>0.475</td>
<td>0.370</td>
<td>0.110</td>
<td>0.035</td>
<td>-</td>
<td>1.500</td>
<td>0.250</td>
<td>6.1</td>
<td>0.3</td>
<td>0.25</td>
<td>0.230</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Available Water Capacity** $\theta_A = (\theta_{\text{FC}} - \theta_{\text{PWP}})$ [-] 

33
The experiments can be divided into two sets of simulations with distinct objectives: The first set compares the impact of the choice between the two soil maps, while the second set investigates the involved physical processes by testing the role of individual soil parameters. These two sets of simulations are described in two separate sections.

Comparison of impact of FAO and JRC soil maps

The standard soil map used in COSMO-CLM is derived from the Digital Soil Map of the World (FAO, 2003), which is available at a 5 arc minutes resolution and in geographical projection. It is based on the FAO map published in 1974 (FAO/UNESCO, 1974). A more recent and better resolved product, the Soil Geographical Database of Eurasia (see Lambert et al., 2002), which is part of the European Soil Database, was released by the European Commission Joint Research Center (JRC) in 2006 (European Commission and the European Soil Bureau Network, 2004); it contains soil texture data over Europe at a 1-km resolution. More details about the conversion of the FAO and JRC products into TERRA_ML-compatible maps for use in simulations are given in Appendix A.2.

In a first set of simulations we aim at investigating the sensitivity of regional climate simulations to the choice of the applied soil map. In a first simulation (FAO) we used the standard FAO map and in a second simulation (JRC) we replaced this standard map with the soil map from the JRC product. Figure 2.1 displays both soil maps as used in the two simulations. These differ in some regions, with coarser/finer soil textures in JRC compared to FAO as shown in Fig. 2.2. Note that the values of the soil parameters that are attributed to each soil class are shown in Table 2.1 and correspond to the
2.2. EXPERIMENTAL SETUP

Figure 2.2: Change in soil textural class expressed as JRC-FAO. Red (blue) colours indicate finer (coarser) soil grains in JRC compared to FAO. Grid points with special soil classes (“rock”, “ice”, “peat”) on either map and which exhibit a change in soil class appear in grey.

standard values in TERRA_ML.

Role of the different soil parameters

To help understand the physical processes underlying the climate response to the modified soil class associated with the respective soil maps, we perform additional experiments which isolate the effect of specific soil parameters. Although in each of these simulations the FAO soil map was used, soil class conversions (changes from one soil class to another) were introduced by modifying specific soil parameters for one soil class at a time.

We selected the following three representative soil class conversions (frequently occurring when changing the soil map from FAO to JRC):

- sandy loam to loam (experiments “SL2L”)
- loam to loamy clay (experiments “L2LC”)
- loam to clay (experiments “L2C”)

For each of these conversions we examined the individual influence of the following sets of soil parameter:

- Field capacity $\theta_{FC}$ and permanent wilting point $\theta_{PWP}$ (experiments “WHC”)
- Hydraulic conductivity parameters $K_0$ and $K_1$ (experiments “COND”)
### Table 2.2: Summary of the simulations. The standard look-up table is shown in Table 2.1, while other look-up tables are the same except for 2 parameters in one soil class.

- Hydraulic diffusivity parameters $D_0$ and $D_1$ (experiments “DIFF”)

Only hydrological parameters have been chosen since they cause most of the climate effect, as will be shown in Section 2.3.1. Note that, as described in the Appendix A.1.2, Hydraulic conductivity and hydraulic diffusivity are physically linked. Our experiments, however, allow us to disentangle the effect of gravity (COND) versus capillary forces (DIFF) on the vertical water transport.

In TERRA\_ML, the hydraulic diffusivity $D_w$ is defined as a function of the soil liquid water content $\theta$, and $D_0$ and $D_1$ are constants in this function (see Equation A.18 in Appendix A.1.2). As a result, modifying $D_0$ and $D_1$ together changes the sensitivity of hydraulic diffusivity to soil moisture and, since these two parameters do not appear anywhere else in the model it seems reasonable to combine them together for our present purpose.

Since hydraulic conductivity $K_w$ is defined similarly to hydraulic diffusivity (see Equation A.19), $K_0$ and $K_1$ were also grouped together in a set of parameters. We note that, in addition to influencing $K_w$, $K_0$ also plays a role in the parameterization of bare soil evaporation (see Equations A.5 and A.7).

Finally, the field capacity $\theta_{FC}$ and permanent wilting point $\theta_{PWP}$ are selected and modified together because of their role in controlling the amount of water available for plant transpiration (see Equation A.12 in Appendix A.1). On the one hand, $\theta_{FC}$ is the amount of water that remains in the soil after excess water has drained out. On the other hand, $\theta_{PWP}$ is the minimum amount of water necessary to prevent plants from wilting and below which almost no transpiration takes place anymore.
2.2. **EXPERIMENTAL SETUP**

Figure 2.3: Relative changes (expressed for $X$ as $\frac{X_{JRC} - X_{FAO}}{|X_{FAO}|}$) in selected soil parameters for 3 selected soil class conversions. Absolute values of each parameter for each soil class are given in Table 2.1.

Note that, for several reasons, neither porosity (volume of void, $\theta_{PV}$) nor air dryness point ($\theta_{ADP}$) were modified in our experiments. First, we want to isolate the effect of field capacity $\theta_{FC}$ and permanent wilting point $\theta_{PWP}$, which are expected to be most crucial for evapotranspiration (see Equation A.12). In addition, changes in $D_w$ and $K_w$ could arise from modifying $\theta_{PV}$ and $\theta_{ADP}$ (see Equations A.18 and A.19), while modifying $\theta_{FC}$ and $\theta_{PWP}$ does not directly impact hydraulic diffusivity and conductivity, thus isolating more strictly the various influences.

The name of each experiment reflects the set of modified soil parameters and the involved soil class conversion. In total, combining three soil class conversions with three sets of parameters leads to nine additional simulations. All simulations are summarized in Table 2.2. For instance, for testing the impact of $\theta_{FC}$ and $\theta_{PWP}$ when modifying the soil class from “sandy loam” (in the FAO soil map) to “loam” (in the JRC soil map), the values of $\theta_{FC}$ and $\theta_{PWP}$ of the soil class “sandy loam” were replaced by those of “loam”. This simulation is called WHC-SL2L. Note that the values of all other parameters are kept as in Table 2.1 for this simulation. Similarly, in two other simulations called COND-SL2L and DIFF-SL2L the hydraulic conductivity and diffusivity parameters, respectively, of “sandy loam” were set to the values of “loam”. Thus, at the points where the soil class conversion SL2L occurs, the impact of the tested parameters on climate can be compared to the impact of changing the soil map, i.e. the impact of changing all parameters together.

A limitation of these simulations is that the changes affect all grid points with the corresponding soil class, while in JRC only some of these grid points are affected. In addition, grid points with other soil classes are not modified. Thus, these simulations are not entirely comparable to the full JRC experiment. However, the impact of these two differences is likely to be very restricted since the changes that we focus on in our analysis are mostly local (as shown in Section 2.3.1). Therefore, we assume that these small differences
do not impact our results in a significant way.

Relative changes in selected soil parameter values for the three selected soil class conversions are displayed in Fig. 2.3. Note that although all the selected soil class conversions lead to a finer soil texture, not all parameters change in a similar way. As expected, both field capacity and permanent wilting point increase; by contrast, the resulting available water capacity \( \theta_A = (\theta_{FC} - \theta_{PWP}) \), which represents the amount of water potentially available for plants, can either increase or decrease. Similarly, hydraulic diffusivity and conductivity parameters can change in either direction. However, since the final values of hydraulic conductivity and diffusivity depend on soil moisture as well, it is difficult to assess the overall change in these two parameters.

2.3 Results

Section 2.3.1 presents results from the comparison between simulations with both tested soil maps (FAO and JRC), while Section 2.3.2 presents results from additional simulations that investigate the role individual soil parameters.

2.3.1 Impact of the new JRC soil map (JRC versus FAO)

Mean climate and surface fluxes

Since the differences between simulations FAO and JRC are largest in summer (not shown), we concentrate on this period (June-August, JJA) for the whole analysis. Figures 2.4(a,b) display changes in summer mean 2-meter temperature and precipitation between the two simulations (JRC minus FAO). The region north of the Black Sea experiences warmer (up to 2K), drier (up to 0.5 mm/day) summer with the JRC soil map. On the other hand, the Baltic region (mainly Poland and Belarus) and the region over Italy and the Western Balkan states experience cooler (up to 1K), wetter (up to 0.4 mm/day) summer with this soil map. As shown in Table 2.3, changes in temperature clearly depend on the soil class conversion, while changes in precipitation are slightly less related to the soil class conversion pattern and thus more difficult to interpret.
2.3. RESULTS

Figure 2.4: Difference between JRC and FAO (JRC-FAO) for mean summer climate (JJA). Sensible and latent heat fluxes are computed only over land and at points where the respective mean value is positive in both simulations, while evaporative fraction (i) is computed only at points where both turbulent fluxes are positive (gray shading indicates location where this is not the case). Note that a line smoothing has been applied for display purposes.
### Table 2.3: Spatial mean and standard deviation of changes in climate variables (JJA mean) for the main soil class conversions (JRC-FAO). Soil classes are “sand” (S), “loam” (L), “clay” (C) and intermediate classes “sandy loam” (SL) and “loamy clay” (LC), while “all” refers to all soil classes from S to C, plus the special class “peat”. Soil class conversions are named after the soil class in FAO and the soil class in JRC, separated by “2”.

<table>
<thead>
<tr>
<th>Soil class in FAO</th>
<th>SL</th>
<th>L</th>
<th>L</th>
<th>SL</th>
<th>L</th>
<th>LC</th>
<th>all classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil class in JRC</td>
<td>L</td>
<td>LC</td>
<td>C</td>
<td>S</td>
<td>S</td>
<td>L</td>
<td>as in FAO</td>
</tr>
<tr>
<td>Soil class conversion</td>
<td>SL2L</td>
<td>L2LC</td>
<td>L2C</td>
<td>SL2S</td>
<td>L2S</td>
<td>LC2L</td>
<td>none</td>
</tr>
<tr>
<td>Number of grid points</td>
<td>555</td>
<td>513</td>
<td>236</td>
<td>493</td>
<td>166</td>
<td>158</td>
<td>4411</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>t2m (°C)</th>
<th>Precip. (mm/day)</th>
<th>H (W/m²)</th>
<th>λE (W/m²)</th>
<th>SW\textsubscript{net} (W/m²)</th>
<th>LW\textsubscript{net} (W/m²)</th>
<th>VWC (%)</th>
<th>SMI (⁻)</th>
<th>EF (⁻)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.21±0.26</td>
<td>0.01±0.20</td>
<td>-5.6±7.0</td>
<td>7.5±7.7</td>
<td>-0.35±2.09</td>
<td>1.13±1.25</td>
<td>7.9±2.4</td>
<td>0.11±0.18</td>
<td>0.066±0.076</td>
</tr>
<tr>
<td></td>
<td>0.73±0.49</td>
<td>-0.21±0.20</td>
<td>12.9±6.0</td>
<td>-15.2±7.1</td>
<td>2.76±2.47</td>
<td>-4.26±2.82</td>
<td>2.9±2.0</td>
<td>-0.14±0.08</td>
<td>-0.135±0.061</td>
</tr>
<tr>
<td></td>
<td>1.23±0.58</td>
<td>-0.18±0.22</td>
<td>24.7±8.6</td>
<td>-30.0±10.5</td>
<td>3.11±2.27</td>
<td>-6.59±2.90</td>
<td>9.2±2.9</td>
<td>-0.23±0.12</td>
<td>-0.255±0.085</td>
</tr>
<tr>
<td></td>
<td>-0.21±0.23</td>
<td>0.03±0.16</td>
<td>6.9±6.0</td>
<td>6.1±6.3</td>
<td>-0.37±1.81</td>
<td>1.74±1.24</td>
<td>-4.3±2.5</td>
<td>0.12±0.15</td>
<td>0.073±0.066</td>
</tr>
<tr>
<td></td>
<td>0.01±0.32</td>
<td>-0.02±0.17</td>
<td>0.3±4.4</td>
<td>-2.0±5.2</td>
<td>-2.45±2.61</td>
<td>0.30±1.86</td>
<td>-11.5±1.4</td>
<td>-0.02±0.07</td>
<td>-0.003±0.048</td>
</tr>
<tr>
<td></td>
<td>-0.41±0.30</td>
<td>0.04±0.11</td>
<td>-0.3±4.4</td>
<td>-10.5±5.8</td>
<td>-1.34±1.41</td>
<td>2.36±1.78</td>
<td>2.35±1.9</td>
<td>0.09±0.06</td>
<td>0.095±0.050</td>
</tr>
<tr>
<td></td>
<td>0±0.10</td>
<td>-0.01±0.11</td>
<td>0.1±1.2</td>
<td>0.1±1.2</td>
<td>0.00±1.23</td>
<td>0.01±0.74</td>
<td>0.01±0.4</td>
<td>0.00±0.02</td>
<td>-0.001±0.011</td>
</tr>
</tbody>
</table>
To put the impact of soil types on the mean climate in the broader context of uncertainties in RCMs, Fig. 2.5(a) displays changes in mean 2m-temperature for selected regions as compared to the PRUDENCE model intercomparison (Jacob et al., 2007). We show the PRUDENCE inter-model interquartile range (referred to as PRUDENCE IQR, positive by definition and plotted from $-\text{IQR}/2$ to $+\text{IQR}/2$), which gives a measure of the spread among different RCMs for a given region. The overall effect of changing the soil map is small compared to the PRUDENCE IQR and, mostly, not significant, due to offsetting effects from different soil type conversions within a given region. On the other hand, when considering only a single soil type conversion, changes are often large and of similar magnitude as the PRUDENCE IQR, although always smaller. However, since none of the PRUDENCE regions covers the area exhibiting the strongest effect on climate (i.e. North of the Black Sea; only the Eastern European (EA) region covers part of it, but also includes a large area of cooling over Poland and Belarus), we added for comparison a region defined between 30 and 45 degrees East and 45 and 52 degrees North (BS). Although PRUDENCE IQR over BS is not available, the impact is striking, with the change in mean temperature over the region due to the soil map being as large as the inter-model IQR from other regions. This shows that, in some regions, soil type specifications can lead to differences in mean summer climate as large as typical differences between RCMs.

The mechanisms controlling these differences are associated to changes in the hydrological properties of the soil. Changes in sensible ($H$) and latent ($\lambda E$) heat fluxes are large (up to more than 30 W/m$^2$, see Fig. 2.4g,h and Table 2.3) and, although they also mostly compensate each other (their sum is about 3-5 W/m$^2$, i.e. of the same order as net radiation), their properties explain temperature changes quite well. More specifically, the changes in these fluxes ($H$ and $\lambda E$) correspond to a modification of surface energy partitioning, expressed by the evaporative fraction $\text{EF} = \lambda E / (H + \lambda E)$ shown in Fig. 2.4(i). In regions with increased EF, the part of the available energy at the surface which is used for evapotranspiration (E, expressed as $\lambda E$ in energy units) increases (i.e. $\lambda E$ increases and $H$ decreases, given that the available energy ($\simeq H + \lambda E$) remains approximately constant). This change in partitioning is confirmed by Fig. 2.7(a), where changes in $H$ and $\lambda E$ are plotted for the main soil class conversions: Changes in $H$ and $\lambda E$ are of similar magnitude and opposite sign. Therefore, EF is a good indicator of the changes in both fluxes. Since $H$ directly influences air temperature, 2m-temperature increases in regions of decreased EF and decreases in regions of increased EF, respectively.

By contrast, radiative properties do not strongly affect the local climate in our simulations. First, changes in net shortwave and longwave radiation are in most cases smaller than changes in turbulent fluxes (Table 2.3). In
Figure 2.5: Changes in summer 2m-temperature (a) mean and (b) interannual variability (standard deviation of JJA means) for selected regions. “JRC-FAO” refers to all land cells within the region, while “no change” refers to grid cells with the same soil type in the two simulations. Colored bars refer to grid points corresponding to soil type conversions. Error bars show the upper and lower quartiles of changes at individual grid cells and numbers indicate the number of grid cells available for each bars. “PRUDENCE IQR” refers to the interquartile range of models means from the PRUDENCE model intercomparison experiment over the corresponding region, computed with the mean value of each model and plotted from $-\text{IQR}/2$ to $+\text{IQR}/2$. Values for PRUDENCE IQR are derived from Table 3 in Jacob et al. (2007). Iberian Peninsula (IP), Mid-Europe (ME), Scandinavia (SC), Mediterranaen (MD) and Eastern Europe (EA) are the largest regions as defined by Christensen and Christensen (2007) and an additional region North of the Black Sea (BS) is defined as 30 to 45°E, 45 to 52°N. Note that for the additional region BS, data from PRUDENCE is not available.
addition, Figures 2.4(d,e), which show net longwave and shortwave radiation, emphasize that although there is a non-negligible change in both radiation fluxes (up to 10 W/m$^2$), they mostly compensate each other (note the inverse scale). Thus, total net radiation only differs by about 3-5 W/m$^2$ (not shown). Furthermore, changes in net shortwave radiation are driven by changes in incoming direct shortwave radiation (solar radiation) due to changes in cloud cover (Fig. 2.4f), while changes in net longwave radiation result from changes in outgoing longwave radiation due to temperature changes (since longwave radiation emission is a function of temperature). We also note that changes in albedo are small for most soil class changes (see Table 2.1), which consolidates our interpretation.

In addition, like other studies that have shown that surface fluxes driven by soil moisture can influence atmospheric circulation (e.g. Fischer et al., 2007a), we note slight changes in mean sea level pressure in our simulations. However, these changes are very small and they cannot explain the identified major changes in temperature and precipitation. Indeed, as shown in Table 2.3, the effects are local and grid points where the soil class is the same in the two simulations (i.e. conversion “none”) do not exhibit any change in climate compared to other points.

**Soil moisture**

Surface energy partitioning in transitional climate regions usually depends on soil moisture (SM) since this variable controls $\lambda E$ and thus EF there (Seneviratne et al., 2010). SM can be expressed by different metrics. Two of them are used here. First, the Volumetric Water Content (VWC) expresses SM as a volumetric fraction, i.e. $VWC = \frac{\text{volume of water in } V}{V}$ where $V$ is a soil volume. Second, Soil Moisture Index (SMI) expresses SM relative to field capacity $\theta_{FC}$ and plant wilting point $\theta_{PWP}$ as $SMI = \frac{\theta - \theta_{PWP}}{\theta_{FC} - \theta_{PWP}}$ where $\theta$ is SM expressed as VWC. In other words, SMI describes the amount of water within the available water capacity $\theta_A$ and therefore the water stress, with no stress for $SMI = 1$ and no water available for $SMI = 0$ (Betts, 2004; Seneviratne et al., 2010). The layers considered in this Section for both VWC and SMI cover the root depth, thus capturing the water stress for the plants and therefore transpiration (see Equation A.11).

As shown in Fig. 2.6(a), changes in VWC over the root depth cannot explain changes in EF. Indeed, VWC increases over most regions, including north of the Black Sea and over the Baltic region, and decreases over other regions where EF increases (e.g. over Italy and the Western Balkan states). By contrast, SMI does explain these changes very well (Fig. 2.6b), with regions where EF and SMI increase (e.g. Baltic region, Italy and the Western Balkan states) and other regions where both EF and SMI decrease (e.g. north of the Black Sea). This better correspondance of SMI to EF, compared to that of VWC to EF, is due to the changes in soil parameters
Figure 2.6: Change in summer (JJA) soil moisture within the rootzone (JRC-FAO), expressed as two different measures: (a) Volumetric Water Content (VWC) [m]; (b) Soil Moisture Index (SMI) [-]. Note that a line smoothing has been applied for display purposes.

(mainly $\theta_{FC}$ and $\theta_{PWP}$) which modify the sensitivity of E to VWC, while SMI accounts for these changes (especially for transpiration, through Equations A.11 and A.12). This is in line with previous studies (see Seneviratne et al., 2010) showing that EF is better related to relative rather than absolute soil moisture content. Figure 2.7(b) shows relative changes in these two variables for the main soil class conversions and confirms a relationship between them, although there is a large spread in the response of EF to changes in SMI.

This spread can be explained when comparing the respective maps of the two variables (SMI and EF; Figures 2.6b and 2.4i, respectively). In spite of a good visual agreement in sign between the changes in SMI and EF (as well as $H$ and $\lambda E$), the intensities of these changes differ, with large changes in the two turbulent fluxes in southern Europe, where changes in SMI tend to be rather small. This explains most of the spread in Fig. 2.7(b) and it simply reflects the different evapotranspiration regimes. In the South, EF is limited by SMI and, therefore, even small SMI changes impact EF. In the North, EF tends to be rather radiation-limited; there, SMI does not play an important role for the local climate. Thus, changes in EF are largest in the South. By contrast, although the latitudinal gradient in changes in SMI is less marked, it reflects the negative feedback loop between EF and SM: in the South, an increase in SM leads to an increase in EF, which then depletes SM, thus damping the initial SM increase. By contrast, in the North, an increase in SM does not strongly impact EF and this negative feedback loop does not exist.
2.3. RESULTS

(a) Latent heat vs sensible heat flux

(b) Change in EF vs change in root zone SMI

(c) Change in EF vs latitude

(d) Legend: grid points per soil class conversion

Figure 2.7: Plots of changes in mean summer values for selected variables. Each point corresponds to one grid point, its colors indicating the soil class conversion (from FAO to JRC), as indicated in the legend (d).
This explains why changes in SMI are largest in the North and relatively small in the South. Figure 2.7(c) summarizes these findings; it shows the changes in EF with latitude for the main soil class conversions. There is indeed a tendency toward smaller changes in EF at high latitudes to some extent, although only two soil class conversions are present at these latitudes (SL2S and SL2L, in black and red, respectively) and thus it is difficult to generalize this statement. We note that these two soil class conversions are also those that show the largest spread in Fig. 2.7(b), thus providing further support for our interpretation that the level of agreement between changes in EF and in SMI is affected by latitude. At latitudes lower than about 55 to 60°N, no clear gradient can be identified.

Overall, changes towards a finer soil texture tend to lead to lower SMI, which in turn induces lower EF and thus leads to a warmer, mostly drier climate. The impact on temperature appears clearly, while the impact on precipitation remains more patchy.

Soil moisture-precipitation feedback

Patterns of changes in precipitation are not as well related to soil class conversion patterns than patterns of changes in temperature. In most cases, regions of increased EF correspond to regions of increased precipitation, but there are exceptions. Although this behaviour is rather indicative of a positive soil moisture-precipitation coupling, there are also a few regions with an indication of negative soil moisture-precipitation coupling and/or of non-local effects. This confirms other studies, which have highlighted the possibility of both positive and negative coupling depending on the conditions and the location (see e.g. Seneviratne et al., 2010, for a review). Nonetheless, it should be noted that the Tiedtke convection scheme used in this model version was found to mostly lead to positive soil moisture-precipitation feedback in a study for the alpine region (Hohenegger et al., 2009). As a final remark, one should consider that the final change in soil moisture level in the simulations results both from the change in soil class and from the modified precipitation.

Interannual variability

In addition to having an impact on the mean climate, the change in soil map also affects its interannual variability (IAV). Figure 2.5(b) shows changes in interannual variability for 2m-temperature, expressed as the standard deviation of summer means and compares it to inter-model interquartile range from the PRUDENCE experiment, in the same way Fig. 2.5(a) shows it for the mean. Like for the mean, the total changes over each region (JRC-FAO) is generally smaller than the PRUDENCE IQR, although local effects can be large over some soil type conversions. Note that, for a given soil type conversion, IAV can increase or decrease depending on the region, while the
2.3. RESULTS

The sign of the effect on the mean is consistent throughout regions. The largest effect on IAV occurs over Scandinavia, where the mean is only marginally affected. This results from the increase occurring for all soil type conversions in this region. Conversely, the effect over BS is small, unlike for the mean; in particular, the impact of the soil type conversion L2C is spread around 0.

Overall, changes in soil types affect interannual climate variability, but the reasons and the processes involved do not seem to be directly related to specific soil type conversions. A given soil type conversion can lead to both an increase or a decrease in IAV depending on the region, and the behaviour is therefore difficult to predict.

2.3.2 Role of the individual soil parameters

Among the six soil class conversions applied to the largest number of grid points (i.e. those displayed in Table 2.3), we consider three conversions in Section 2.3.2 (SL2L, L2LC, L2C). To some extent, LC2L is also covered since it is the opposite of L2LC, although we cannot exclude hysteresis effects. The two remaining conversions are not selected for further investigations: L2S exhibits very small changes in climate, while SL2S mainly concerns grid points far north of the continent (over e.g. Finland and Norway).

Changes in VWC, SMI and EF for each experiment in reference to FAO are shown in Fig. 2.8 as a function of latitude and are contrasted for the three soil class conversions. The impact of the individual parameter sets (DIFF, WHC, COND) is compared to the overall impact of the soil class conversion (JRC versus FAO).

For the three soil class conversions, changes in VWC mostly depend on $\theta_{FC}$ and $\theta_{PWP}$, as expected since these parameters modify the amount of water that can be stored in the soil. However, as highlighted in Section 2.3.1, VWC is not a good indicator of the impact on climate; therefore, SMI is also displayed in Figures 2.8(b,e,h). Changes in SMI are due to changes in two sets of parameters: $\theta_{FC}$ and $\theta_{PWP}$, on the one hand, and hydraulic diffusivity parameters ($D_0$ and $D_1$) on the other hand, while hydraulic conductivity parameters ($K_0$ and $K_1$) are not found to play a substantial role. Finally, changes in EF (Fig. 2.8c,f,i) confirm this result: modifying hydraulic conductivity parameters alone does not substantially impact EF, implying a minor role of these parameters in our simulations; by contrast, both other sets of parameters (hydraulic diffusivity parameters; field capacity and plant wilting point) have a strong impact on EF. More specifically, for two soil class conversions (SL2L and L2C) field capacity $\theta_{FC}$ and plant wilting point $\theta_{PWP}$ explain most of the total changes in EF while hydraulic diffusivity parameters ($D_0$ and $D_1$) also appears to contribute substantially to this change. The remaining soil class conversion (L2LC) shows the opposite behavior, with
(a) Change in VWC, SL2L
(b) Change in SMI, SL2L
(c) EF vs latitude, SL2L

(d) Change in VWC, L2LC
(e) Change in SMI, L2LC
(f) EF vs latitude, L2LC

(g) Change in VWC, L2C
(h) Change in SMI, L2C
(i) EF vs latitude, L2C

**Figure 2.8:** Changes in VWC (left), SMI (center) and EF (right), with respect to FAO, versus latitude for JRC and the simulations with individual sets of parameters modified, for the three selected soil class conversion: sandy loam to loam (SL2L, top), loam to loamy clay (L2LC, middle), and loam to clay (L2C, bottom). VWC and SMI are computed over the root zone. Only the points for which JRC has the corresponding soil class conversion are shown. Lines show a running mean for bins of 2° in latitude.
θ\text{FC} and θ\text{PWP} explaining a minor part of EF changes and \( D_0 \) and \( D_1 \) contributing to it more strongly.

Thus, these two sets of parameters impact the amount of moisture available for evapotranspiration and, thereby, EF. Interestingly, changes in EF are not always very well related to changes in SMI, despite our findings from Fig. 2.7. For instance, for the soil class conversion L2LC, changes in SMI at low latitudes are about the same for WHC-L2LC and DIFF-L2LC, while changes in EF differ substantially. This hints at other controls on EF than only water availability; for instance, the vertical transport of water within the soil may have been affected in a different way for these two simulations, which might in turn have affected the distribution of water within the soil layer and therefore EF in some situations, ultimately.

In all three cases, an increase in both field capacity and plant wilting point together (\( \theta\text{FC} \) and \( \theta\text{PWP} \)) leads to higher VWC, as expected since more water can be stored into the soil. Higher water available capacity \( \theta_A \) leads to higher SMI and EF, but a closer comparison of the soil class conversions (i) L2C and (ii) L2LC shows that this might be linked to other variables. First of all, and as indicated in Fig. 2.3, the change in \( \theta_A \) is largest in L2C, although the values of \( \theta\text{FC} \) and \( \theta\text{PWP} \) change less than in L2LC. By contrast, the corresponding changes in EF (and SMI) are much larger in L2LC, which suggest that \( \theta_A \) is not necessarily representative of changes in \( \theta\text{FC} \) and \( \theta\text{PWP} \). Changes in both field capacity \( \theta\text{FC} \) and plant wilting point \( \theta\text{PWP} \) are larger in L2LC and could explain this behaviour; however, indirect effects do probably play an important role as well. In particular, we note that larger changes in VWC in L2LC strongly impact hydraulic diffusivity (\( D_w \), see Equation A.18) which then further impacts SMI and its relationship to EF. Therefore, we conclude that both the available water capacity \( \theta_A \) and the absolute values of \( \theta\text{FC} \) and \( \theta\text{PWP} \) play a substantial role.

Increasing hydraulic diffusivity \( D_w \) leads to an increase in VWC. First, Equation A.18 shows that \( D_w \) increases with increasing \( D_0 \), \( D_1 \) and VWC (\( \theta \)), and that there is a feedback between VWC and \( D_w \) since they influence each other. In the conversion SL2L, both \( D_0 \) and \( D_1 \) increase, leading to an enhanced hydraulic diffusivity. This increased diffusivity further enhances the VWC and forms a positive feedback loop. Similarly, in the conversion L2LC, a the decrease in both \( D_0 \) and \( D_1 \) is correlated with a decrease in VWC. For the last conversion (L2C), results are more difficult to interpret given that changes in \( D_0 \) and \( D_1 \) are of opposite sign and their relative importance can hardly be assessed due to the highly non-linear relationship between these two variables.

Interestingly, the vertical transport of water within the soil due to capillary forces appears to be a driving factor in our experiments. Indeed, the transport of water between the soil layers, which is expressed by Equation A.16, has two components. First gravitational drainage, expressed by
K_w, is not critical in our simulations, since we find very little sensitivity of soil moisture and EF on K_0 and K_1. Second, capillary forces, represented by D_w and the vertical gradient of water within the soil, play an important role as shown in simulations where D_0 and D_1 are modified. In addition to this direct effect, changes in D_w through changes in VWC (Equation A.18) lead to a similar but indirect effect. Although it is difficult to disentangle this indirect effect from other possible effects due to changes in VWC, vertical profiles of soil moisture (Fig. 2.9) support this hypothesis: The experiments DIFF exhibit the vertical profile of SMI closest to JRC, except for the soil class conversion SL2L where changes in hydraulic diffusivity parameters (in particular D_0) were small and other parameters dominate the observed effects. Vertical profiles of VWC show that, while experiments with modified WHC are closest to JRC in terms of mean and absolute value, the DIFF experiments exhibit the correct profile shape. In all cases, the COND experiments correspond exactly to the original FAO simulation, emphasizing the negligible role of hydraulic conductivity. This analysis highlights the fact that, in addition to SMI, the vertical redistribution of water within soil layers strongly impacts EF and the climate. In our experiments, the critical variable for this vertical distribution is hydraulic diffusivity.

The focus of this study is on the overall effect on the total evapotranspiration. Since transpiration is the main component of evapotranspiration in our simulations, it dominates our results. Changes in bare soil evaporation (not shown) were substantially different in all cases and they can help understand the underlying mechanisms. For instance, while in JRC this component was not much different from its value in FAO, in WHC-SL2L it differed substantially from FAO. In WHC-L2LC, bare soil evaporation even increased while it decreased in JRC. Both hydraulic conductivity and hydraulic diffusivity parameters did not play any substantial role for bare soil evaporation, and these results suggest that the parameters that control this component of E are quite different from those controlling transpiration.

Among the investigated parameters, the hydraulic diffusivity parameters (D_0 and D_1) and field capacity \( \theta_{FC} \) and plant wilting point \( \theta_{PWP} \) explained most of the EF response in our soil class experiments. Our choice of parameters looks quite reasonable, since we reproduce this response to a large extent.

2.4 Discussion and conclusions

We performed and analyzed RCM simulations with different soil maps and soil parameters over Europe for the period 1980-2005. These experiments highlight the important role of soil parameters for summer climate, for the most part via their impact on EF, especially in regions where soil moisture is a limiting factor for evapotranspiration. This impacts the mean summer
2.4. DISCUSSION AND CONCLUSIONS

Figure 2.9: Mean vertical profiles of volumetric water content (VWC) and soil moisture index (SMI) in summer for the three selected soil class conversions: sandy loam to loam (SL2L, left), loam to loamy clay (L2LC, middle), and loam to clay (L2C, right). Analyses are done only for the points for which JRC has the corresponding soil class conversion compared to FAO. All simulations are displayed for each conversion.
climate by up to about 2°C for temperature and 20% for precipitation over regions with large differences between soil texture datasets, while the impact on interannual climate variability, which is found to be more difficult to relate to changes in soil type, appears to be smaller, except for the increase over Scandinavia. Comparison with the multi-model analysis from PRUDENCE (Jacob et al., 2007) reveals that, over most regions, changes in mean summer 2-meter temperature are small compared to inter-model interquartile range due to compensating changes within regions. However, over some regions such as North of the Black Sea, changes are as large as the PRUDENCE inter-model interquartile ranges of other regions (a direct comparison is not possible because PRUDENCE regions do not include this region). This shows that the choice of a soil dataset can potentially have an impact on the mean summer climate in some regions as large as the choice of the RCM itself.

More specifically, in simulations where individual parameters were modified, we identify the important role of parameters affecting the available water capacity (field capacity \( \theta_{FC} \) and permanent wilting point \( \theta_{PWP} \)) as well as hydraulic diffusivity (parameters \( D_0 \) and \( D_1 \)). In particular, the impact of \( \theta_{FC} \) and \( \theta_{PWP} \) on values of VWC and, therefore, on hydraulic diffusivity highlight the fact that, although model-specific range of VWC might not be problematic for the parameterization of E if SMI is represented appropriately, the dynamic of soil moisture and its vertical profile is influenced by hydraulic diffusivity and, therefore, by absolute values. In other words, a model formulated in a similar way similar as in COSMO-CLM, i.e. with E parameterized mainly as a function of SMI and with hydraulic diffusivity/-conductivity expressed as a function of VWC, needs correct values for both these soil moisture variables in order to model the processes in a correct way over time. Distinguishing strictly between the effect of field capacity and plant wilting point versus hydraulic diffusivity parameters, and their interactions, would have been interesting but is difficult, precisely because they are intimately related to one another. This could be done using methods of factor separation as described by e.g. Stein and Alpert (1993), but additional simulations would have been required. Although we could not clearly distinguish between the effect of these two sets of parameters for these reasons, our results indicate an especially important role of \( D_w \).

The results show a negligible sensitivity of soil moisture dynamics and profile to hydraulic conductivity, in contrast to the strong sensitivity to hydraulic diffusivity. Here, we recall that these two variables describe a single property of the soil, namely the ability of water to flow within it, but they express this property in different units (see Appendix A.1.2). In the current formulation of the land-surface scheme TERRA_ML, \( K_w \) represents the gravity term while \( D_w \) represents capillary forces. Physically, these two parameters are intimately linked; therefore, modifying them independently is not fully realistic, but allows us to distinguish between the effect of gravitational
drainage (through $K_w$) and capillary movement (through $D_w$). As expected, gravity only plays a marginal role in summer since the water content is almost always kept below field capacity. By contrast, capillary forces play an important role for the vertical motion of water within the soil. Note that, in some land-surface models (e.g. Community Land Model, see Lawrence et al., 2011), hydraulic conductivity accounts for the effects of both gravity and capillary forces. In these models, hydraulic conductivity is likely to play a key role, in a similar way as parameters controlling hydraulic diffusivity in TERRA ML do.

The amplitude of the differences in mean summer climate for different soil classes (up to about $2^\circ C$ in 2-meter temperature and 20% in precipitation) provides evidence that the soil class plays an important role for the local climate in summer. Since the JRC soil map is assumed to be more accurate and up-to-date, one might expect improvements in the simulated climate with this new soil map, even at a coarse resolution. However, we note that COSMO-CLM does not perform better with the new JRC soil map. For instance, spatial root mean square error for summer mean 2m-temperature is larger in the region North of the black sea or over Poland (not shown). Several reasons may have contributed to this result. First of all, physical improvements in models do not necessarily lead to a reduction of bias, given the necessary preexisting model tuning. This is particularly true given the structure of the model and its heritage from BATS, which implies that retuning would probably be necessary. Second, the conversion of the original soil database from the JRC into TERRA ML classes and the desired spatial resolution is subject to uncertainties, and the method used differs from the one used for the FAO soil map, due to different soil categories in the two original datasets (see Appendix A.2). Finally, the values of the different soil parameters are critical, and these are subject to uncertainties as well. That said, this study focuses on the physical processes that lead to the simulated differences and not on improving model performance.

The region investigated in our study covers the European domain. We note that, in other regions, the processes involved might be different. In particular, other types of soils such as organic soils might play an important role, as shown by e.g. Lawrence and Slater (2008), who identified that the effect of soil carbon content on thermal and hydrological soil properties can lead to changes of about $2.5^\circ C$ in mean summer 2-meter temperature.

In this study, we only investigate the effect on the mean climate and its interannual variability. However, since soil moisture is crucial for extremes events such as heat waves and droughts (e.g. Seneviratne et al., 2006b; Lorenz et al., 2010; Hirschi et al., 2011), soil parameters are expected to strongly impact these events. The representation of these effects would benefit from improved soil databases as well, which is crucial given their relevance for society. In addition, given the impact of soil properties on soil moisture dynamics
and therefore on its memory and its impact on climate, not only climate simulations but also weather and seasonal forecasts would benefit from consistent databases of soil properties. Major issues include the discrepancy between models in the range and effect of parameters, the heterogeneity of soil parameters values in space and the variability of these parameters within a given soil class, which often exceeds their variability between classes.

Although much research has focused on the impact of land-use changes and related vegetation properties and their interactions with climate, which can even be investigated in details by LSMs that include dynamic vegetation, nothing comparable has been undertaken for soils. Given the large impact of soil specification on climate, soil classes could also be developed and included in a more dynamic way, at least for long-term climate simulations. Indeed, some soil physical properties can change depending on e.g. crops, crop management, land clearing and land use (Uhland, 1950; Ghuman et al., 1991; Alegre and Cassel, 1996; Zimmermann et al., 2006). In addition, interactions between soil and vegetation may play a role as well, as suggested by Osborne et al. (2004). For instance, the organic content of the soil can change relatively quickly after deforestation. Conversely, soil properties may subsequently impact vegetation and its development by providing conditions that favour certain species.

Finally, soil parameters being highly relevant in transitional regions between dry and wet climate given their relation to soil moisture dynamics and surface fluxes, disagreements between models with respect to land-climate interactions might be in part linked to these soil parameters as well (for instance land-atmosphere coupling, see Koster et al., 2004). Hence our results highlight the need to characterize soil class parameters in better detail in land surface and climate models.

Acknowledgments

We thank Alessandro Dosio (JRC) for information about the European Soil Database. The European Soil Database (ESDB) has been developed by the Land Management & Natural Hazards Unit of the European Commission Joint Research Centre, in the context of the development of European Environmental Data Centres by the European Commission and the European Environment Agency.
Land-surface controls on afternoon precipitation diagnosed from observational data: uncertainties and confounding factors

Atmos. Chem. Phys., 14, 8343–8367, doi:10.5194/acp-14-8343-2014 *
Benoit P. Guillod¹, Boris Orlowsky¹, Diego Miralles²,³, Adriaan J. Teuling⁴, Peter D. Blanken⁵, Nina Buchmann⁶, Philippe Ciais⁷, Michael Ek⁸, Kirsten L. Findell⁹, Pierre Gentine¹⁰, Benjamin R. Lintner¹¹, Rus-

*This publication was slightly changed from its original version by adapting abbreviations to ensure consistency throughout this thesis.
CHAPTER 3. LAND–PRECIPITATION COUPLING

sell L. Scott\textsuperscript{12}, Bart Van den Hurk\textsuperscript{13}, and Sonia I. Seneviratne\textsuperscript{1}

**Abstract** The feedback between soil moisture and precipitation has long been a topic of interest due to its potential for improving weather and seasonal forecasts. The generally proposed mechanism assumes a control of soil moisture on precipitation via the partitioning of the surface turbulent heat fluxes, as assessed via the evaporative fraction (EF), i.e., the ratio of latent heat to the sum of latent and sensible heat, in particular under convective conditions. Our study investigates the poorly understood link between EF and precipitation by relating the before-noon EF to the frequency of afternoon precipitation over the contiguous US, through statistical analyses of multiple EF and precipitation data sets. We analyze remote-sensing data products (Global Land Evaporation: the Amsterdam Methodology (GLEAM) for EF, and radar precipitation from the NEXt generation weather RADar system (NEXRAD)), FLUXNET station data, and the North American Regional Reanalysis (NARR). Data sets agree on a region of positive relationship between EF and precipitation occurrence in the southwestern US. However, a region of strong positive relationship over the eastern US in NARR cannot be confirmed with observation-derived estimates (GLEAM, NEXRAD and FLUXNET). The GLEAM–NEXRAD data set combination indicates a region of positive EF–precipitation relationship in the central US. These disagreements emphasize large uncertainties in the EF data. Further analyses highlight that much of these EF–precipitation relationships could be explained by precipitation persistence alone, and it is unclear whether EF has an additional role in triggering afternoon precipitation. This also highlights the difficulties in isolating a land impact on precipitation. Regional analyses point to contrasting mechanisms over different regions. Over the eastern US, our analyses suggest that the EF–precipitation relationship in NARR is either atmospherically controlled (from precipitation persistence

\textsuperscript{1}Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland.
\textsuperscript{2}Ghent University, Laboratory of Hydrology and Water Management, Ghent, Belgium.
\textsuperscript{3}University of Bristol, School of Geographical Sciences, Bristol, UK.
\textsuperscript{4}Wageningen University, Hydrology and Quantitative Water Management Group, Wageningen, the Netherlands.
\textsuperscript{5}University of Colorado, Department of Geography, Boulder, CO, USA.
\textsuperscript{6}ETH Zurich, Institute of Agricultural Sciences, Zurich, Switzerland.
\textsuperscript{7}Laboratoire des Sciences du Climat et de l’Environnement, LSCE, Gif-sur-Yvette, France.
\textsuperscript{8}National Centers for Environmental Prediction, Suitland, MD, USA.
\textsuperscript{9}Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA.
\textsuperscript{10}Columbia University, Department of Earth and Environmental Engineering, New York, NY, USA.
\textsuperscript{11}Rutgers, The State University of New Jersey, Department of Environmental Sciences, New Brunswick, NJ, USA.
\textsuperscript{12}Southwest Watershed Research Center, USDA-ARS, Tucson, AZ, USA.
\textsuperscript{13}Royal Netherlands Meteorological Institute, KNMI, De Bilt, the Netherlands.
and potential evaporation) or driven by vegetation interception rather than soil moisture. Although this aligns well with the high forest cover and the wet regime of that region, the role of interception evaporation is likely overestimated because of low nighttime evaporation in NARR. Over the central and southwestern US, the EF–precipitation relationship is additionally linked to soil moisture variations, owing to the soil-moisture–limited climate regime.

3.1 Introduction

Soil-moisture–precipitation feedback has been investigated for several decades and, despite some progress in recent years, remains a poorly understood process and a large source of uncertainty in climate models (Seneviratne et al., 2010). While studies until the 1990s tended to focus on the concept of moisture recycling (i.e., the fraction of precipitation directly contributed by regional evaporation from the land surface; see Seneviratne et al., 2010), more recent studies have emphasized the importance of indirect feedback mechanisms – that is, an influence of soil moisture on atmospheric stability, boundary layer characteristics, and thereby precipitation formation (e.g., Schär et al., 1999; Pal and Eltahir, 2001; Findell and Eltahir, 2003a; Ek and Holtslag, 2004; Betts, 2004; Santanello et al., 2009; Hohenegger et al., 2009; Taylor et al., 2011; Lintner et al., 2013; Gentine et al., 2013). Such indirect effects can theoretically lead to feedbacks of either sign (Seneviratne et al., 2010). For instance, over wet soils, humidity input into the boundary layer increases, but turbulence and boundary layer height decrease; the interplay of these two effects with the environment can trigger or suppress convective rainfall locally depending on the prevailing conditions (e.g., Ek and Holtslag, 2004; Gentine et al., 2013). Although most studies report a positive feedback, some suggest the existence of a negative feedback in certain regions (Findell and Eltahir, 2003a,b; Cook et al., 2006; Hohenegger et al., 2009; Westra et al., 2012; Gentine et al., 2013). Furthermore, nonlocal processes can also be important (e.g., Taylor and Ellis, 2006). In particular, spatial heterogeneity of soil moisture has been shown to possibly induce mesoscale circulations favoring precipitation over dry soils, for example in the Sahel region (Taylor et al., 2011) but also globally (Taylor et al., 2012).

The entire soil-moisture–precipitation feedback can be decomposed into a chain of processes as follows (Fig. 3.1, modified from Seneviratne et al., 2010; see also, e.g., Santanello et al., 2011):

A. Soil moisture impacts the partitioning of energy at the land surface into sensible and latent heat flux (\(H\) and \(\lambda E\), respectively), as quantified by the evaporative fraction \(\text{EF} = \frac{\lambda E}{H + \lambda E}\).

B. The moisture and heat input to the atmosphere corresponding to changes in EF impacts subsequent precipitation.
Figure 3.1: Schematic description of the soil-moisture–precipitation coupling and feedback loop. Positive arrows (blue) indicate processes leading to a positive soil-moisture–precipitation feedback (wetting for positive soil moisture anomaly, drying for negative soil moisture anomaly), the negative arrow (red) indicates a potential negative feedback damping the original soil moisture anomaly, and the red–blue arrow indicates the existence of both positive and negative feedbacks between evaporative fraction (EF) and precipitation anomalies. (A), (B), and (C) refer to the different steps of the feedback loop (see text). Modified from Seneviratne et al. (2010).

C. Precipitation impacts soil moisture by replenishing the soil moisture reservoir.

Relationship A (higher soil moisture leading to higher EF) is expected to be most significant in regions that are transitional between wet and dry climates, where soil moisture is the main limiting factor for land evaporation (e.g., Koster et al., 2004; Seneviratne et al., 2006b; Teuling et al., 2009a; Hirschi et al., 2011). Note here the potentially negative feedback within relationship A (red arrow in Fig. 3.1), since increased soil moisture content enabling high evaporation leads to faster depletion of the soil moisture, thus dampening the initial evaporation increase (see also Seneviratne et al., 2010; Boé, 2013). Relationship B, i.e., higher EF leading to higher (or lower) precipitation, is generally the most uncertain part of the soil-moisture–precipitation coupling and feedback and can exhibit positive or negative sign through boundary layer regulation (e.g., Ek and Holtslag, 2004; Santanello et al., 2007; van Heerwaarden et al., 2009). The impact of precipitation on soil moisture (relationship C), on the other hand, can be considered as straightforward, albeit with a dependence on the partitioning of precipitation into interception, runoff, and infiltration. Some studies investigate single relationships (e.g., relationship A; see for instance Dirmeyer, 2011), while A–B has been analyzed as one relationship (e.g., Taylor et al., 2012) as well as by combining metrics from each individual relationship (A and B; e.g., Dirmeyer et al., 2012).
3.1. INTRODUCTION

The existence, the sign, and the strength of soil-moisture–precipitation coupling, i.e., the impact of soil moisture on precipitation (relationship A–B), and in particular EF–precipitation coupling (B), remain heavily debated in the literature.

Modeling studies yield contrasting results, identifying both positive (Schär et al., 1999; Pal and Eltahir, 2001; Koster et al., 2004) and negative soil-moisture–precipitation relationships in some cases (Findell and Eltahir, 2003a,b; Ek and Holtslag, 2004; Hohenegger et al., 2009; Siqueira et al., 2009; van den Hurk and van Meijgaard, 2010). It has been shown that model-based studies suffer from deficiencies, such as the dependence on the chosen convective parameterization or resolution (e.g., Hohenegger et al., 2009). In particular, Taylor et al. (2013) suggest that current convective parameterizations in models lead to a positive feedback in regions where observations and cloud-resolving simulations indicate negative feedback. Dirmeyer et al. (2006) highlight large biases in global climate models (GCMs) with respect to covariability between key atmospheric and land-surface variables, and Koster et al. (2003) suggest that soil-moisture–precipitation feedbacks may be overestimated in GCMs.

Given the large range of results from modeling studies, observational studies are necessary. However, for a number of reasons, these have been largely inconclusive (Seneviratne et al., 2010). First, the scarcity of soil moisture and EF measurements is a recurrent limitation. In particular, while recent satellite remote-sensing efforts have facilitated global analyses and generated new insights (e.g., Taylor et al., 2012), these only provide data of soil moisture in the top few millimeters of the soil and in regions without dense vegetation cover. This is often not representative of deeper layers and, thus, of EF, especially in vegetated areas. Second, we note that one of the most challenging tasks in assessing soil-moisture–precipitation coupling (i.e., A–B) from observational data is to establish causal rather than mere statistical links between soil moisture (or EF) and precipitation (see also Salvucci et al., 2002; Orlowsky and Seneviratne, 2010).

The difficulty of causal inferences from observational data arises from two main confounding effects. First, given the influence of precipitation on soil moisture (process C) it can be difficult to assess whether a detected relationship between soil moisture and precipitation is due to A–B, C, or both. In particular, persistence in precipitation at various timescales (from synoptic to interannual scales, including seasonal scale) can induce apparent causal links, for which even lagged correlations, such as between soil moisture and subsequent precipitation, may in fact simply reflect relationship C. Second, covariability between two variables (for instance soil moisture and convective precipitation) may be a necessary but not a sufficient condition for a causal link since it does not exclude the possibility that both quantities are governed by a third influencing variable (for instance sea surface temperature;
see Orlowsky and Seneviratne, 2010). Ideally, potential confounding variables should be taken into account in observational analyses; this is, however, rarely done in practice, mostly due to difficulties in identifying confounding variables or lack of data availability.

In order to overcome the issue of data scarcity, some studies have used state-of-the-art reanalysis products (e.g., Bisselink and Dolman, 2008; Findell et al., 2011). Soil moisture and associated land-surface fluxes in reanalysis products are, however, ultimately model-based and therefore share the deficiencies of their land-surface models. Some reanalysis products assimilate screen-level variables (temperature, humidity) in order to better constrain the surface energy budget (Mahfouf, 1991; Bouttier et al., 1993b,a; Gentine et al., 2011b) and may thus be advantageous over other reanalysis products. Nonetheless, such land data assimilation procedures may introduce biases in surface variables (e.g., Betts et al., 2003; Seneviratne et al., 2004). In addition, reanalyses suffer from other issues such as the lack of mass conservation. Finally, they suffer from the similar difficulties in isolating causal relationships as the studies based on observational data, although they provide a more comprehensive data basis. Therefore, reanalysis-based investigations are a useful complement to, but ultimately cannot replace, observational studies.

In this study, we investigate soil-moisture–EF–precipitation coupling (i.e., processes A and B, with a focus on B) over North America, addressing the aforementioned issues. We use direct observations of EF and precipitation from FLUXNET sites, remote-sensing-derived products (satellite-driven EF estimates from GLEAM and precipitation from the US radar network NEXRAD), and soil moisture, EF, and precipitation from the North American Regional Reanalysis, NARR (see Sec. 3.2). Specifically, we quantify the relationship between before-noon EF (and soil moisture) and afternoon convective rainfall occurrence via the triggering feedback strength (TFS; see Findell et al., 2011, and Sec. 3.3). This metric suggests, when applied to NARR, a region of positive coupling over the eastern US (Findell et al., 2011). Here, we first compare TFS estimates derived from observation-driven data sets with those from NARR (Sec. 3.4). We then consider the potentially confounding role of precipitation persistence on TFS (Sec. 3.5), and further investigate the role of soil moisture and vegetation interception storage on land evaporation, as well as the inferred EF–precipitation coupling (Sec. 3.6). Finally, results from these sections and their implications are discussed in Sec. 3.7.

### 3.2 Data sets

We provide here a description of the data sets used in this study. The analysis is restricted to North America for consistency with Findell et al. (2011). The data sets include a reanalysis product (the North American Regional
Reanalysis, hereafter referred to as NARR), ground-based point-scale observations from FLUXNET, and remote-sensing-derived products: the NEXt generation weather RA
dar system (NEXRAD) and Global Land Evaporation: the Amsterdam Methodology (GLEAM). For 3-hourly data sets (NARR and GLEAM), the 3h UTC time step closest to each local 3h time period (in standard local time based on longitude) is used, as in Findell et al. (2011). Thus, a lead or lag of up to 1h may occur between the data sets.

3.2.1 NARR

The North American Regional Reanalysis (NARR; see Mesinger et al., 2006) is maintained at the National Center for Environmental Prediction (NCEP) and spans the period from 1979 to present. With its high spatial (about 32km horizontal) and temporal (3h) resolution, it allows for analyses focused on the diurnal evolution of land–atmosphere variables, which is an important aspect when analyzing the impact of surface fluxes on convection and precipitation. Its key characteristic is that it successfully assimilates high-quality precipitation observations into the atmospheric analysis, contrary to other reanalyses. This might in principle allow for a more realistic representation of land hydrology and land–atmosphere interactions. Humidity observations are also assimilated to constrain the atmospheric state, but they do not directly constrain surface fluxes via soil moisture nudging. Some other variables that affect the land surface, such as screen-level temperature, are not assimilated (Mesinger et al., 2006). Surface radiation fluxes can also be significantly biased in NARR (Kennedy et al., 2011). Moreover, West et al. (2007) identified spurious grid-scale precipitation events and related them to anomalous latent heating in cases of strong mismatch between assimilated and modeled precipitation. Ruane (2010a,b) highlighted that, while the exaggerated model precipitation is reduced by the assimilation of precipitation observations, other components of the water cycle such as evaporation and moisture convergence are not corrected. Indeed, assimilation products do not conserve water.

The land component of NARR is the Noah land-surface model (Ek et al., 2003). The soil includes four layers spanning the following depths: 0–10cm, 10–40cm, 40cm–1 m, and 1–2 m. Bare soil evaporation (plant transpiration) is limited by soil moisture in the top layer (root zone), and evaporation from vegetation interception is accounted for. The root zone is defined for each grid cell as a function of vegetation type – at the analyzed sites, it includes the top three or four layers.

Here, we use NARR data from the years 1995–2007, and most of the analyses are restricted to days when data are available from other data sets (NEXRAD and GLEAM; see Secs. 3.2.3 and 3.2.4, respectively). This removes possible impacts of different time periods or time series lengths. Analyses of the longer 1979–2007 period are included in the Appendix B (Secs. B.1
and B.4) for comparison, yielding similar results.

All data are adjusted to local time by taking the 3 h period closest to the standard local time. Thus, for afternoon values, for instance, (12–6 p.m.), data from 09:00–15:00 UTC are used west from 247.5° E while 06:00–12:00 UTC data are used for the rest of the continent.

### 3.2.2 FLUXNET

FLUXNET is a global network of micrometeorological measurement sites (Baldocchi et al., 2001; Baldocchi, 2008), which uses the eddy-covariance method to measure exchanges of CO$_2$, water and energy between the land surface and the atmosphere. It currently includes over 500 sites worldwide (http://www.fluxnet.ornl.gov/introduction) with a relatively large density over Europe and North America. The density of the network as well as the record lengths in these regions allow for spatial analyses. FLUXNET is the largest available network of “direct” observations of latent and sensible heat fluxes, which, in spite of some known issues (underestimation of the fluxes and lack of energy balance closure, point-scale measurements with relatively small footprint area, possible change in footprint depending on, for example, wind direction), provides largely model-independent data and is therefore a direct estimate pertinent to our analyses.

In this study, we use data from the FLUXNET LaThuile dataset, a global standardized database of eddy covariance measurements which includes a large number of sites. Measurements of sensible ($H$) and latent ($\lambda E$) heat fluxes are used to compute EF, while global radiation (i.e., incoming shortwave, $R_g$) and potential global radiation (i.e. extraterrestrial radiation, $R_{g}^{\text{pot}}$) are used to get a proxy for cloud cover (see Sec. 3.3.2). One of the main issues with eddy-covariance measurements is that the energy balance is not closed: the sum of $H$ and $\lambda E$ typically underestimates the available energy by 10–30% (e.g., Wilson et al., 2002; Mauder et al., 2006; Foken, 2008; Hendricks Franssen et al., 2010). However, as we do not use $H$ and $\lambda E$ directly but only through EF, we note that the commonly used “fixed Bowen ratio” correction for the energy balance closure (i.e. attributing the missing energy to latent and sensible heat fluxes while keeping the Bowen ratio $B_w = \frac{H}{\lambda E}$ constant, e.g. Blanken et al., 1997) does not affect EF. Hence, we can expect that EF is only marginally affected by the lack of energy closure at the sites.

A total of 39 sites, listed in Table 3.1, are used in this study, all of them located in the US and Canada. The selection of the sites is based on several criteria: first, coverage by precipitation radars from NEXRAD (see Sec. 3.2.3) as well as $R_g$ measurements are requirements for use in our study. Second, summers with many gaps in any of the required variables are removed, and only sites with a reasonable amount of remaining data are kept for the analysis ($\gtrsim 100$ days).
3.2. DATA SETS

3.2.3 NEXRAD

The NEXt generation weather RADar system (NEXRAD) is a network of 159 Weather Surveillance Radar-1988 Doppler (WSR-88D) sites covering the United States. Data are archived at the National Climatic Data Center (NCDC) of the US National Weather Service. Here, we use the one-hour precipitation product (N1P) from the level III data. More details about NEXRAD products can be found at [http://www.ncdc.noaa.gov/oa/radar/radarresources.html](http://www.ncdc.noaa.gov/oa/radar/radarresources.html) (accessed on 20 December 2012). N1P data for summer (June to August, JJA) from 1995 to 2007 were downloaded at NEXRAD stations covering FLUXNET sites and their vicinity. We use 3 hr averages of precipitation within 20 km around each FLUXNET site. Aggregating with different radii and time-averaging methods leads to robust results (not shown).

3.2.4 GLEAM

GLEAM (Global Land Evaporation: the Amsterdam Methodology; see Miralles et al., 2011b) is a global data set of daily land-surface evaporation ($E$) based on satellite observations, available at a resolution of 0.25°. Estimates of $E$ for day $i$ are derived from

$$E_i = E_{i}^{\text{pot}} S_i + (1 - \beta) E_{I_i},$$  \hspace{1cm} (3.1)

where $E_{i}^{\text{pot}}$ is the potential evaporation (at day $i$), derived through the Priestley and Taylor formulation (Priestley and Taylor, 1972) using data of net radiation ($R_{\text{net}}$) and near-surface air temperature. $S_i$ denotes the evaporative stress (at day $i$) and is computed combining (a) observations of vegetation water content (microwave vegetation optical depth) and (b) estimates of root-zone soil moisture ($\theta_i$) from a multilayer soil module driven by observations of precipitation ($P_i$) and surface soil moisture ($\theta_{\text{obs}}$). The inclusion of vegetation optical depth accounts for the effects of plant phenology; its low day-to-day variability causes minor effects on the short-term dynamics of $E_i$. $E_{I_i}$ denotes the vegetation rainfall interception loss, calculated based on Gash’s analytical model of rainfall interception (Gash, 1979) and described in Miralles et al. (2010); $\beta$ is a constant to account for declines in transpiration when the canopy is wet (see Miralles et al., 2010, 2011b).

The satellite-data-driven evaporation model GLEAM is based on a larger array of satellite information than other evaporation products, which often apply algorithms requiring variables that are difficult to retrieve from satellite data (e.g., near-surface humidity and wind speed), and therefore rely on reanalysis forcing. To our knowledge, GLEAM is also the only large-scale satellite-data-driven evaporation product that estimates the temporal dynamics of root-zone soil moisture (based on observations of precipitation and surface soil moisture and a multilayer soil model). This root-zone soil moisture...
<table>
<thead>
<tr>
<th>Site</th>
<th>Lat [°N]</th>
<th>Lon [°E]</th>
<th>Altitude [m]</th>
<th>IGBP class</th>
<th>Years available</th>
<th>Years excluded</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA-Mer</td>
<td>45.41</td>
<td>−75.52</td>
<td>70</td>
<td>WET</td>
<td>1998–2005</td>
<td>2000</td>
<td>Roulet et al. (2007)</td>
</tr>
<tr>
<td>US-Bkg</td>
<td>44.35</td>
<td>−96.84</td>
<td>510</td>
<td>GRA</td>
<td>2004–2006</td>
<td></td>
<td>Saito et al. (2009)</td>
</tr>
<tr>
<td>US-Ho1</td>
<td>45.20</td>
<td>−68.74</td>
<td>60</td>
<td>ENF</td>
<td>1996–2004</td>
<td></td>
<td>Fernandez et al. (1993)</td>
</tr>
<tr>
<td>US-Ho2</td>
<td>45.21</td>
<td>−68.75</td>
<td>91</td>
<td>ENF</td>
<td>1999–2004</td>
<td></td>
<td>Fernandez et al. (1993)</td>
</tr>
</tbody>
</table>

Table 3.1: FLUXNET sites included in this study, with latitude, longitude, altitude, vegetation class (IGBP, International Geosphere Biosphere Programme), years available, years excluded from the analysis and reference publication. IGBP classes represented in this subset of sites are: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forests (DBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), permanent wetlands (WET) and woody savannas (WSA). For a detailed description of the vegetation classes, see [http://www.fluxdata.org/DataInfo/default.aspx/](http://www.fluxdata.org/DataInfo/default.aspx/), accessed on 21 June 2013. This table continues on the next page.
<table>
<thead>
<tr>
<th>Site</th>
<th>Lat</th>
<th>Lon</th>
<th>Altitude</th>
<th>IGBP</th>
<th>Years available</th>
<th>Years excluded</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-SP3</td>
<td>29.75</td>
<td>−82.16</td>
<td>50</td>
<td>ENF</td>
<td>1999–2004</td>
<td>1999</td>
<td>Bracho et al. (2011)</td>
</tr>
<tr>
<td>US-Ton</td>
<td>38.43</td>
<td>−120.97</td>
<td>177</td>
<td>WSA</td>
<td>2001–2006</td>
<td></td>
<td>Ma et al. (2007)</td>
</tr>
</tbody>
</table>

**Table 3.1:** FLUXNET sites included in this study (cntd.).
is used to constrain the atmospheric demand for water calculated based on radiation and temperature (note that explicit soil moisture constraints are not directly included in analogous models; e.g., Su, 2002; Mu et al., 2007; Fisher et al., 2008). GLEAM estimates of $E$ have been extensively validated and compared to other methodologies (Miralles et al., 2011a,b; Mueller et al., 2013; Liu et al., 2013; Trambauer et al., 2014; Miralles et al., 2014a,b). In particular, GLEAM was successfully validated using measurements from 163 eddy-covariance stations and 701 soil moisture sensors all across the world and run with a wide range of data sets for the required input variables in Miralles et al. (2014b). $AE$ estimates from GLEAM have been applied to a large number of studies over the past 3 years (e.g., Miralles et al., 2011a, 2012, 2014a,b; Reichle et al., 2011; Mueller et al., 2013; Liu et al., 2013; Fersch and Kunstmann, 2013; Jasechko et al., 2013; Trambauer et al., 2014), and the estimates error has been characterized using triple collocation (Miralles et al., 2011a).

We use a version of GLEAM that is driven by the input data sets noted in Table 3.2. Importantly, precipitation from NEXRAD (see Sec. 3.2.3) is used as input (to estimate interception loss and drive the soil module). GLEAM estimates using three other precipitation data sets yield similar results (Appendix B.2, Fig. B.3). GLEAM usually operates at daily time steps; as shown in Eq. (3.1), the computation of $E$ requires daily estimates of potential evaporation, $E_{i}^{\text{pot}}$; evaporative stress, $S_{i}$; and interception, $E_{I,i}$. Here, to estimate before-noon EF (9 a.m.–12 p.m., i.e., $EF_{i,9-12}$), several modifications to the original methodology are therefore necessary.

GLEAM is first run with daily input variables aggregated to days beginning/ending at around 9 a.m. standard local time. The resulting estimates of root-zone soil moisture ($\theta_{i-1}$) used to derive $S_{i-1}$ roughly correspond to 9 a.m. on day $i$, as they are derived using the cumulative precipitation up to 9 a.m. and instantaneous observations of surface soil moisture from the early morning hours (between 1.30 a.m. and 6 a.m. depending on the satellite platform – see Table 3.2 for details on the soil moisture remote-sensing products). In the assimilation, early morning surface soil moisture observations are combined with the bucket model estimates of soil moisture based on the rainfall until 9 a.m. Although surface soil moisture might not always be representative of root-zone soil moisture, Miralles et al. (2014b) found mild improvements in the root-zone soil moisture estimates of GLEAM after assimilating the satellite observations.

Before-noon EF at day $i$ (i.e., $EF_{i,9-12}$) is then computed using $S_{i-1}$ estimates as a proxy for the before-noon evaporative stress conditions. Since days with morning-time precipitation are not included in the computations of the TFS, $E_{I,i,9-12}$ is assumed to be zero. $EF_{i,9-12}$ is therefore calculated
### Table 3.2: Data sets used in GLEAM product.

Daily aggregates are computed locally to match the before-noon Evaporative Fraction (EF) estimate (i.e., starting and ending at around 9 a.m., see Sec. 3.2.4). $S$ and $E^{\text{pot}}$ are the evaporative stress and the potential evaporation, respectively. See Sec. 3.2.4 for details.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Data set</th>
<th>Resolution and use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil moisture</td>
<td>NASA-LPRM (Owe et al., 2008) based on SSMI (1995–2002) and AMSR-E (from mid-2002)</td>
<td>night-time overpass (for the $S$ calculation)</td>
</tr>
<tr>
<td>Vegetation optical depth</td>
<td>NASA-LPRM (Owe et al., 2008)</td>
<td>daily (for the $S$ calculation)</td>
</tr>
<tr>
<td>Precipitation</td>
<td>NEXRAD (Sec. 3.2.3)</td>
<td>daily (for the $S$ calculation)</td>
</tr>
<tr>
<td>Net radiation</td>
<td>GEWEX SRB 3.0 (Stackhouse et al., 2004)</td>
<td>daily (for the $S$ calculations) and 3 hourly frequencies (for the morning $E^{\text{pot}}$)</td>
</tr>
<tr>
<td>Air temperature</td>
<td>NCEP-1 (Sheffield et al., 2006)</td>
<td>daily (for the $S$ calculations) and 3 hourly frequencies (for the morning $E^{\text{pot}}$)</td>
</tr>
</tbody>
</table>
as

\[
EF_{i,9-12} = \frac{\lambda E_{i,9-12}^\text{pot} S_{i-1}}{R_{i,9-12}^\text{net} - G_{i,9-12}},
\]

where \(R^\text{net}\) is net radiation from the GEWEX SRB data set (satellite-based product; see Stackhouse et al., 2004) and \(G\) is the ground heat flux, computed as a function of \(R^\text{net}\) and land cover type according to Miralles et al. (2011b). Note that the focus on the inter-day rather than intra-day variability in \(EF\) is advantageous since \(EF\) is considered most stable around noontime (e.g., Gentine et al., 2007). In addition, \(EF\) is rather robust to \(G\) given the low day-to-day variability in \(G\) relative to its diurnal cycle.

To summarize, \(EF_{i,9-12}\) is computed in two steps:

1. GLEAM is first run as in Miralles et al. (2011b) to derive the daily averages of evaporation (\(E_i\)) and evaporative stress (\(S_i\))—see Eq. (3.1). The only difference here is that we compute daily values from about 9 a.m.–9 a.m. for all variables (depending on longitude but always before 9 a.m.).

2. \(S_{i-1}\) is used to calculate before-noon \(EF\) (i.e., \(EF_{i,9-12}\)) using Eq. (3.2).

In this form, \(S_{i-1}\) accounts for evaporative stress due to soil moisture deficits only and does not account for interception. This is done to acknowledge that interception rates are high even at night (see e.g., Pearce et al., 1980) and therefore vegetation only remains wet for a few hours after rainfall (4 ± 1.9 h using values from field studies compiled by Miralles et al., 2010), and because days with morning rainfall are removed, not being the subject of our analyses (Sec. 3.3.2).

Nonetheless, to allow comparison with NARR, we introduce an alternative formulation which accounts for interception evaporation during the before-noon time period by assuming that vegetation stores intercepted water from the previous-day precipitation. To do so, we use a modified stress formulation, \(S^*_i\), which assumes that water remains on vegetation from the previous-day precipitation. Hence, we can rearrange Eq. (3.1) as

\[
E_i = E_i^\text{pot} S_i + (1 - \beta) E_{I,i} = E_i^\text{pot} S^*_i,
\]

which yields

\[
S^*_i = S_{i-1} + (1 - \beta) \frac{E_{I,i-1}}{E_i^\text{pot}}.
\]

Estimates of \(EF_{i,9-12}\) can then be computed using \(S^*_{i-1}\) instead of \(S_{i-1}\) in Eq. (3.2) to account for interception evaporation. Nonetheless, this alternative approach is likely unrealistic due to the above-mentioned fast evaporation rates. In addition, findings from field studies highlight that advection and downward sensible heat flux rather than radiation are critical to the evaporation of intercepted water (e.g., Pearce et al., 1980; Asdak et al., 1998;
3.3. METHODS

Holwerda et al., 2012), and therefore the contribution of interception evaporation to the (radiation-based) EF is not straightforward. Nonetheless, we use this alternative approach as a sensitivity test of potential interception effects in Sec. 3.6 (see Fig. 3.11).

Note that the timing of the input data sets for the $S$ and $S^*$ computation is crucial to this application, in particular for precipitation. First, we do not want to include any information about afternoon precipitation for the estimated before-noon EF on the same day. Second, rainfall occurring in the night preceding the estimated EF must be included in order to get an EF reflecting the conditions in the early morning. Unfortunately, the definition of “days” in many standard daily precipitation products varies, as shown in Table B.1 in the Appendix B, and is sometimes unclear: for instance, the use of data from the Global Precipitation Climatology Project (GPCP; see Huffman et al., 2001) is inappropriate due to the time window of the data set (00:00 to 00:00 UTC, i.e., from 4 p.m. (7 p.m.) to 4 p.m. (7 p.m.) in the US west (east) coast; see Table S1 in the Supplement). Also noteworthy, for the CPC Unified gauge product (Chen et al., 2008) days are defined differently depending on the country. For most of the US, the defined window is 12:00 to 12:00 (UTC, i.e., 4 a.m.–4 a.m. in the west coast/7 a.m.–7 a.m. in the east coast), which in principle suits our requirements, although uncertainties remain due to differing reporting times between contributing rain gauge stations. NEXRAD is not affected by this issue given its higher temporal resolution.

Due to the large diversity of precipitation products and the sensitivity of EF to precipitation, GLEAM has been driven with several precipitation data sets as input (see discussion in Appendix B.2). Data sets used for this sensitivity test are NEXRAD, CPC Unified (Chen et al., 2008) and PERSIANN (Hsu et al., 1997). These three data sets either suit the required daily time window (like in the case of CPC Unified) or have a subdaily temporal resolution and therefore allow for appropriate daily aggregates (like in the case of NEXRAD and PERSIANN). Results obtained from these three independent precipitation data sets are qualitatively similar (see Fig. B.3 and text in the Appendix B.2).

3.3 Methods

This section describes the convection triggering metric TFS, including the selection of potentially convective days to which the computations are restricted, and the applied statistical test for assessing the significance of the results.
3.3.1 Triggering feedback strength (TFS)

The TFS, defined by Findell et al. (2011), quantifies the link between before-noon EF and afternoon precipitation occurrence as

\[ TFS = \sigma_{EF} \frac{\partial \Gamma(r)}{\partial EF} \]  

(3.4)

where EF is the before-noon evaporative fraction (computed between 9 a.m.–12 p.m. where 12 p.m. is noon), \( \sigma_{EF} \) is the standard deviation of EF and \( \Gamma(r) \) is the probability of afternoon rain (> 1 mm, computed between 12 pm and 6 p.m.). The computation is restricted to summer days (June to August, JJA). Only potentially convective days (Sec. 3.3.2) are included in the computation in order to reduce the impact of large-scale synoptic systems. In addition, surface turbulent fluxes of sensible and latent heat are most likely to impact precipitation formation in convective situations (see Sec. 3.3.2). Note that, like most statistical analyses, a high TFS does not necessarily imply causality between EF and \( \Gamma(r) \) but simply the existence of a statistical correlation between the two variables.

Findell et al. (2011) computed TFS in bins of the parameter space of EF, CTP and HI \(_{\text{low}} \) (the convective triggering potential and a low-level humidity index, respectively; see Findell and Eltahir, 2003a), which are subsequently aggregated. HI \(_{\text{low}} \) is an indicator of humidity in the lower atmosphere, while CTP provides information about atmospheric stability. Accounting for these two variables is expected to reduce possible confounding effects from atmospheric conditions. In our study, however, relatively short observational time series preclude extensive sampling of this parameter space and independent observational sources for CTP and HI \(_{\text{low}} \), i.e. radio soundings, do not exist in the vicinity of all analyzed FLUXNET sites.

We can therefore only approximate the approach of Findell et al. (2011). Thus, we compute here a simplified version of TFS,

\[ TFS^* = \sigma_{EF} \frac{\Gamma(r|EF > EF_{Q60}) - \Gamma(r|EF \leq EF_{Q40})}{EF_{Q80} - EF_{Q20}}, \]  

(3.5)

where \( EF_{QX} \) is the Xth percentile of EF. The variable \( \sigma_{EF} \) and the percentiles of EF are determined for each location and dataset independently. The definition of the bins ensures clearly distinct bins (i.e. no possible overlap even if \( EF_{Q60} = EF_{Q40} \)) while retaining most of the available data. Considering quantiles also partly accounts for different shapes of the EF distributions when comparing different EF datasets. EF values outside of the 0–1 range are excluded from the analysis. Although TFS\(^*\) is an approximation of the original TFS defined by Findell et al. (2011), the two different computations show close agreement when applied to NARR.
3.3. METHODS

3.3.2 Identification of potentially convective days

Midlatitude continental convection tends to occur in the afternoon, as a result of a particular daytime boundary layer evolution (Rio et al., 2009). Potentially convective days are therefore expected to be rain- and cloud-free in the morning. Moreover, convection is usually linked to low atmospheric stability and, therefore, typically positive CTP (Findell et al., 2011).

Findell et al. (2011) therefore identify potentially convective days as days with $\text{CTP} > 0$ and no morning precipitation. In the absence of the necessary information for CTP from observations, we alternatively use the following criteria throughout our analyses:

- No morning precipitation, as in Findell et al. (2011), and
- $R_g/R_g^\text{pot} > 0.67 \max (R_g/R_g^\text{pot})$ in the morning, where $R_g$ is the global radiation (i.e. incoming short-wave) at the land surface and $R_g^\text{pot}$ is the potential $R_g$ in the absence of atmosphere (i.e. extraterrestrial incoming short-wave).

$R_g$ is available from NARR and measured at FLUXNET sites. $R_g^\text{pot}$, being dependent on time and latitude only, is computed for each grid cell used in our analysis for NARR. It is directly available in FLUXNET data. The computation of $\max (R_g/R_g^\text{pot})$, restricted to summer days (JJA), is applied to each site to account for site-specific conditions. $R_g/R_g^\text{pot}$ therefore quantifies the fraction of incoming solar radiation reaching the ground, and its maximum value corresponds to clear-sky cases. Requiring $R_g/R_g^\text{pot} > 0.67 \max (R_g/R_g^\text{pot})$ in the morning is used to remove days with morning clouds from the analysis as they are likely linked to synoptic systems. Cutoff ratios between 0.5 and 0.8 do not lead to different results (not shown). Note that this criterion does not exclude the presence of morning clouds, which could be convective, but prevents cases dominated by morning clouds, likely of stratiform origin.

In this study, the data set combinations use these criteria computed on the following data sets, chosen according to data availability:

- NARR: precipitation and $R_g$ from NARR
- FLUXNET–NEXRAD: precipitation from NEXRAD and $R_g$ from FLUXNET
- GLEAM–NEXRAD: precipitation from NEXRAD and $R_g$ from NARR

The impact of the criteria for the selection of potentially convective days on $\text{TFS}^*$, in particular with respect to the NARR analysis and the different set of criteria used in our study compared to Findell et al. (2011), is small, as discussed in the Appendix B.1 (Fig. B.2).
CHAPTER 3. LAND–PRECIPITATION COUPLING

3.3.3 Statistical tests

The statistical significance of $TFS^* \neq 0$ is tested by bootstrap samples. A $TFS^*$ distribution is computed from 1000 bootstrap samples for which the EF data are kept unchanged and precipitation data are shuffled, which simulates the null hypothesis that no relation between EF and precipitation exists. The bootstrap $TFS^*$ distribution is approximately symmetrical with respect to 0. For a 90% significance level, we require a positive (negative) $TFS^*$ to be at or above (below) the 95 percentile (5 percentile). We chose a rather low significance level of 90% to account for the relatively short time series and the noise inherent in the data.

3.4 TFS from different data sets

The impact of before-noon EF on precipitation occurrence is quantified using the modified triggering feedback strength, $TFS^*$ (see Sec. 3.3). $TFS^*$ is computed at FLUXNET sites from three data set combinations: (i) a reanalysis product (NARR), (ii) direct measurements of surface turbulent heat fluxes at FLUXNET sites for EF in combination with radar precipitation from NEXRAD, and (iii) EF estimates from a satellite-data-driven evaporation product (GLEAM) in combination with NEXRAD precipitation. We compare estimates of $TFS^*$ from these data sets (Sec. 3.4.1) and general characteristics of the EF data sets (Sec. 3.4.2).

3.4.1 TFS patterns

Figure 3.2 displays $TFS^*$ for the three analyzed data set combinations. We note that the NARR pattern reproduces the regions of positive $TFS^*$ from Findell et al. (2011) over the eastern and southwestern US. This shows that our simplified $TFS^*$ computation (Eq. 3.5) reproduces the more sophisticated computation from Findell et al. (2011). Nonetheless, results from Fig. 3.2 (left) display slightly weaker and less significant values, shown in supplementary analyses to be a result of shorter time series (Fig. B.1 in the Appendix). The impact of different sets of criteria for the selection of potentially convective days, another source of discrepancy between our analysis and Findell et al. (2011), turns out to be small (Fig. B.2 in the Appendix).

To complement the maps shown in Fig. 3.2, the distributions of $TFS^*$ values for the three datasets are compared separately over three regions (western, central and eastern US) using box plots (Fig. 3.3). The definition of these regions is based on expected coupling regions from previous studies. The central US region represents an expected soil-moisture–precipitation coupling “hot spot” (e.g. Koster et al., 2004), while the eastern US displays a strong positive EF–precipitation relationship in NARR (Findell et al., 2011). The western US, on the other hand, is a dry region (soil-moisture-limited regime;
3.4. TFS FROM DIFFERENT DATA SETS

![Figure 3.2: Triggering feedback strength (TFS*) in different data sets computed at Fluxnet sites. (left) Evaporative Fraction (EF) and precipitation data from NARR, (center) EF from FLUXNET and precipitation from NEXRAD, and (right) EF from GLEAM and precipitation from NEXRAD. TFS* values significantly different from 0 at the 90% level are indicated by a black asterisk. In case of overlap, points are shifted and the black lines inside the circles indicate the actual location of the station. Empty dots indicate sites with unreliable NEXRAD data.](image)

see Thomas et al., 2009; Schwalm et al., 2012) with little soil moisture and EF variability and is therefore usually not considered conducive to strong soil-moisture–precipitation feedbacks. Strong EF–precipitation coupling is a necessary but not sufficient condition for strong soil-moisture–precipitation coupling.

Generally, FLUXNET displays large variations within each region (Fig. 3.3) and even within smaller climatic regions (e.g. in Florida, Fig. 3.2). It does not confirm the positive TFS* regional pattern evident in NARR over the eastern US (Fig. 3.2). The remote-sensing-derived estimate from GLEAM–NEXRAD displays more consistent patterns, but it also does not yield many significant positive TFS* values in that region (3 significant sites out of 23, Fig. 3.2). Over both the central US and southwestern US, GLEAM–NEXRAD and to some extent FLUXNET show larger TFS* values compared to NARR (Figs. 3.2 and 3.3). We note that inspection of the NEXRAD time series reveals suspect features (not shown) for three sites in the middle of the western region; results with NEXRAD (and GLEAM, which is partly based on NEXRAD) are therefore not shown for these sites (empty dots, e.g., in Fig. 3.3). Results at other sites have been confirmed by analyses with other precipitation data sets (not shown; e.g., with CMORPH; Joyce et al., 2004).

Several reasons might contribute to the observed differences between TFS* estimates from the different data sets, some of which can be discussed with the support of Fig. 3.4 (TFS* for the different combinations of EF and precipitation data sets for the same subset of days, namely the potentially convective days according to the NARR selection):

i. Spatial scale of the EF data: the footprint of FLUXNET measurements is much smaller than the grid cells of NARR and GLEAM (typically 100–2000 m and 25–30 km, respectively; see Sec. 3.4.2). Since
EF–precipitation coupling is expected to occur at scales of about 20–100 km and NEXRAD data are at such a scale, FLUXNET may be less appropriate for this application. Although different TFS* cannot be clearly attributed to differences in footprints based on Fig. 3.4, EF uncertainties are shown to play a strong role in controlling the convection triggering metric (see also Sec. 3.4.2).

ii. Time series length and noise: the lengths of the time series considered here range from a few years in FLUXNET to 13-year (with some gaps) in GLEAM–NEXRAD and NARR. Comparing Fig. 3.2 with the respective panels of Fig. 3.4 shows that the decreased sample size in Fig. 3.4 affects TFS* in NARR and in the GLEAM–NEXRAD combination. A relatively large number of days is required to estimate TFS* robustly, as smaller or noisier samples lead to lower and less significant values. Thus, higher noise levels in observation-based data sets and incomplete sampling due to short record length could explain their weaker values of the metric in the eastern US.
3.4. TFS FROM DIFFERENT DATA SETS

Figure 3.4: Influence of dataset and sample size on TFS*. Only days with data in all datasets are included in the computation, and potentially convective days are further selected based on NARR (see Sec. 3.3.2 for the criteria). TFS* from NARR is boxed in red; TFS* from observation-based combinations in blue. TFS* values significantly different from 0 at the 90% level are indicated by a black star.
iii. Selection of potentially convective days included in the TFS* computation (Sec. 3.3.2): the application of the criteria to different data sets potentially leads to different TFS* estimates, although sensitivity tests do not highlight a strong sensitivity of TFS* to the chosen criteria, as shown in the Appendix for NARR (Fig. ref:chap4:suppl:f:suppl:NARR2).

iv. Other data set characteristics, such as temporal resolution, uncertainties, and possible errors (e.g., modeling components in NARR): such causes for the observed differences are difficult to disentangle from the three above-mentioned factors as the selection of days and the length of the time series are linked to the data sets. While the region of strong EF–precipitation relationship in the eastern US in NARR cannot be confirmed with FLUXNET and GLEAM–NEXRAD, it is possible that time series in these observation-derived data sets are simply too short or too noisy to detect a robust TFS* in this region. Nevertheless, NARR generally exhibits a stronger (weaker) link between EF and convection triggering over the eastern (central and southwestern) US compared to the observation-based estimates used here. Hence our results suggest a product dependence of the derived TFS* patterns.

Hereafter, we focus on the disentangling of these various factors, and in particular on possible fundamental differences in the processes underlying the investigated EF–precipitation relationship. Thereby, analysis of the differences in the data sets themselves might shed light on the different TFS* patterns. Since precipitation data from NARR and NEXRAD agree well in terms of precipitation occurrence (not shown), we focus on the differences between EF data sets and analyze these in the next section.

3.4.2 EF time series

To analyze the agreement of the spatiotemporal dynamics between the three EF data sets, Fig. 3.5 displays their respective correlations with one another in summer (JJA) for before-noon (9 a.m.–12 p.m.) EF. Unlike in the TFS* computation, all days are included in the correlations, but similar results are found for the potentially convective days only. Although positive, correlations are strikingly low at most sites and across all data set combinations. This suggests that the disagreement between the TFS* patterns in the different data set combinations is related to differences in the considered EF data sets (see also Fig. 3.4). Correlations of 10-day and monthly averages of before-noon EF are higher but remain low over the eastern US (Fig. B.4 in the Appendix). Correlations of EF anomalies (i.e., after removing the seasonal cycle within JJA) instead of actual values display similar results (not shown).
3.4. TFS FROM DIFFERENT DATA SETS

Figure 3.5: Correlation of daily JJA before-noon EF values between different data sets. The size of the dots indicates the number of days included in the computation according to the legend shown on the bottom right, and significant correlations at a 99% level are indicated by a black asterisk. Empty dots for GLEAM indicate sites with unreliable NEXRAD (and thus GLEAM) data.

Several reasons might explain these differences. First, the spatial scale over which EF is estimated, or footprint, is data-set-specific, as mentioned above (Sec. 3.4.1, point i). Differences might thus arise from contrasting environmental conditions over the respective footprints (e.g., input of water from rainfall in the case of very local precipitation events), but also from differences in land covers. Indeed, while wet vs. dry periods might be similar in all data sets, some studies have shown that different vegetation might respond differently to given conditions (Teuling et al., 2010). Land cover is in fact different at FLUXNET sites compared to the larger scale in NARR, in particular in regions with cultivated land, as FLUXNET sites are often located over natural vegetation. However, we did not find any systematic link between different land covers and resulting TFS$^*$ (not shown). Similarly, soil texture impacts soil moisture dynamics and EF (e.g., Guillod et al., 2013) and differences in local vs. larger scale soil texture could also be a reason for the differences in EF.

In order to better characterize the EF time series, Fig. 3.6 shows the mean, standard deviation, and persistence (quantified by the decorrelation timescale, $\tau_D$, which integrates the autocorrelation function; see von Storch and Zwiers, 1999) for the three data sets. While we do not find any clear differences between the data sets that can explain the resulting differences in TFS$^*$, the comparison highlights some interesting features. The mean EF is similar in all data sets and exhibits higher values in the eastern US (wetter climate) compared to the drier climate of the western US, although in GLEAM the central US displays even higher mean EF values. The EF standard deviation is noisy, although similar patterns are found across all data sets, with higher EF variability in the central US or in the Southern Great Plains (the exact location depending on the data set). Note, however, that the amplitudes differ widely among the three data sets. This does not necessarily impact TFS$^*$: the change in the probability of afternoon precipitation with respect to EF is scaled by the standard deviation of EF (see Eq. 3.4 and
Figure 3.6: Statistical properties of EF data sets (NARR, FLUXNET, GLEAM, from left to right): (top) mean (EF), (middle) standard deviation (σEF), and (bottom) decorrelation timescale (τd). Only days with data in all three data sets are included in the computation. The decorrelation timescale τd is computed following von Storch and Zwiers (1999). Grey dots indicate too many gaps for a reliable quantification of τd. Empty dots for GLEAM indicate sites with unreliable NEXRAD data.

Berg et al., 2013). Finally, EF persistence is generally lower in the eastern US, suggesting high variability at a scale of one to a few days in this region of strong relationship in NARR (Fig. 3.2, left). Thus, the regions of strong daily correlation between EF and convection triggering correspond, in NARR, to humid regions with low persistence, while in GLEAM–NEXRAD the drier southwestern region, with higher persistence, displays the strongest relationship. For the remaining analyses, we exclude FLUXNET data because of the record length of this data set being too limited.

### 3.5 Impact of EF vs precipitation persistence

Although the TFS metric is a useful tool for investigating the relationship between EF and convective precipitation triggering, precipitation persistence might lead to high TFS even in the absence of an actual impact of EF on precipitation. Here, precipitation persistence refers to precipitation autocorrelation, which might be induced by atmospheric persistence (e.g., from synoptic weather patterns). Resulting persistent wet conditions cause higher EF and vice versa, potentially leading to high TFS* values simply through
3.5. IMPACT OF EF VS PRECIPITATION PERSISTENCE

\[ \Delta \Gamma (EF_{9-12}) \quad \Delta \Gamma (P_{d,prev}) \]

**Figure 3.7:** Difference in the probability of afternoon rainfall on days with high vs. low \( X \), \( \Delta \Gamma (X) \), where \( X \) is the before-noon EF (left panels) or previous-day precipitation (right panels), for NARR (top row) and GLEAM–NEXRAD (bottom row). High (low) \( X \) refers to values higher (lower) than the 60th (40th) percentile of \( X \), i.e., \( \Delta \Gamma (X) = \Gamma (r | X > X_{Q60}) - \Gamma (r | X \leq X_{Q40}) \). Values significantly different from 0 at the 90% level are indicated by a black asterisk. The size of the dots indicates the number of days included in the computation according to the legend shown on the bottom right map. Empty dots indicate sites with unreliable NEXRAD data.

Externally forced precipitation persistence. Although one cannot exclude the possibility that precipitation days cluster together due to a feedback mechanism, atmospheric forcing is a more likely reason. Precipitation persistence might also arise from seasonality in precipitation. However, this effect is less relevant for our study as only summer is considered, and analyses on individual months do not suggest a strong link to seasonality (not shown).

Ideally, the TFS computation should account for such confounding effects, via the filters for potentially convective days (see Sec. 3.3.2). In addition, Findell et al. (2011) use bins of CTP and \( H_{\text{low}} \) in the computation, which we did not implement (Sec. 3.3.1). Nevertheless, we specifically test for the effect of day-to-day precipitation persistence on \( TFS^* \) by replacing before-noon EF with precipitation from the previous day in the \( TFS^* \) computation. With respect to an explanatory variable, \( X \), we denote the change in the probability of afternoon precipitation for high vs. low \( X \) as \( \Delta \Gamma (X) = \Gamma (r | X > X_{Q60}) - \Gamma (r | X \leq X_{Q40}) \). Figure 3.7 (left) shows \( \Delta \Gamma (EF) \) for NARR and the GLEAM–NEXRAD combination, and the patterns strongly resemble those of \( TFS^* \) (Fig. 3.2). Indeed, \( \Delta \Gamma (EF) \) is the term that controls most of the \( TFS^* \) signal, since \( \sigma_{EF} \) and \( \partial \) \( EF \) mostly compensate each other in Eq. 3.4, and maps of \( \sigma_{EF} \) (Fig. 3.6) do not display a pattern similar to that of \( TFS^* \) (Fig. 3.2). Using \( \Delta \Gamma (X) \) allows for a direct comparison between the impact of EF and
that of previous-day precipitation $P_{d,\text{prev}}$, shown on the right of Fig. 3.7 as $\Delta \Gamma(P_{d,\text{prev}})$. In fact, previous-day precipitation is a better predictor for afternoon precipitation occurrence than before-noon EF, which holds for both data sets and across all regions. Given these results, one can wonder whether the signal with EF is, in fact, only reflecting precipitation persistence or whether EF conveys additional information that can help explain afternoon precipitation.

In order to disentangle the impact of EF on precipitation from precipitation persistence, we apply a framework similar to Salvucci et al. (2002) to stratify the data based on previous-day precipitation. Here, only the occurrence of precipitation is considered and we investigate whether the signal emerging with EF reflects previous-day precipitation occurrence alone and thus may be an artifact of precipitation persistence on a short timescale. Note that Salvucci et al. (2002) also accounted for seasonal-scale persistence, which we omit since our analysis is restricted to summer months and our main interest is on short-term persistence (e.g., due to frontal systems or a sequence of these). Figure 3.8 shows $TFS^*$ independent of previous-day precipitation (i.e., as shown before; left column) as well as conditioned on the occurrence of precipitation the day before: here $TFS^*$ is computed for subsets of days with either no precipitation on the previous day or with precipitation on the previous day (center and right columns, respectively). Since the conditioning reduces the number of days available, this analysis is applied to NARR and GLEAM–NEXRAD as well as to the longer set of NARR data, covering 1979–2007 (bottom row) for comparison.

For both NARR and the GLEAM–NEXRAD combination, the signal over the eastern US strongly weakens when days are conditioned on previous-day rainfall (Fig. 3.8). This suggests an important role of precipitation persistence on subsequent precipitation and thus on $TFS^*$. Note, however, that the shorter length of the time series after filtering days based on previous-day precipitation might also impact the results: using all available years from NARR (1979–2007, bottom row), $TFS^*$ is less sensitive to the conditioning on the previous day’s rainfall, where EF might provide information on afternoon precipitation that is additional to previous-day precipitation occurrence. Nonetheless, for days following rain-free days, the weakening of the signal suggests a relevant role of precipitation persistence. Over the southwestern US, the signal appears less sensitive to day-to-day precipitation persistence as $TFS^*$ remains significant at most sites for both data sets.

Overall, precipitation persistence plays an important role and thereby affects $TFS^*$ in all data sets. Several factors can lead to high precipitation persistence via the atmosphere, such as atmospheric dynamics or SST forcing linked with large-scale teleconnection patterns. That said, we cannot exclude a partial contribution of EF–precipitation coupling to the identified persistence features, although larger $\Delta \Gamma$ with previous-day precipitation than with
EF suggests that this is not the dominant mechanism. Finally, the binning in CTP and HI$\text{low}$ might already partly account for this effect in Findell et al. (2011).

3.6 Soil moisture and interception evaporation

In the conceptual framework of a feedback between soil moisture and precipitation via EF (Fig. 3.1), soil moisture is expected to be the main driver of EF. However, our analysis shows that EF can be highly variable from day to day (as reflected, for example, by the low autocorrelation in the eastern US; see Fig. 3.6). This feature is inconsistent with an impact of low-frequency soil moisture variations, which is generally the main relevant factor in the context of weather and seasonal forecasting (e.g., Koster and Suarez, 2001; Seneviratne et al., 2006a; Koster et al., 2010). We thus examine the relevance of soil moisture in the analyzed relationships between land conditions and convection triggering.
We recall that $\lambda E$ (and thereby $\text{EF}$) comprises three main components (Fig. 3.9): plant transpiration ($E_{\text{trans}}$), bare soil evaporation ($E_{\text{soil}}$), and evaporation of water intercepted by vegetation ($E_{t}$). These evaporate water from different reservoirs that typically evolve at different timescales. Root-zone soil moisture ($W_{\text{roots}}$), which reflects precipitation over the previous weeks to months and is affected by vegetation, provides a mid- to long-term storage for $E_{\text{trans}}$. Surface soil moisture ($W_{\text{top}}$, top few centimeters of the soil), which reflects precipitation over the preceding days or week, provides a short-term storage for $E_{\text{soil}}$. Finally, intercepted water on vegetation structures ($W_{\text{canopy}}$), which reflects precipitation over the preceding hours, provides very short storage for $E_{I}$. Although often neglected in climate studies, evaporation of intercepted rainfall has been estimated to represent more than 10% of global terrestrial $E$ (Miralles et al., 2011a) and 20–50% over forests (e.g., Savenije, 2004; McLaren et al., 2008; Gerrits and Savenije, 2011). Typical timescales mentioned here reflect estimates from many studies (see, e.g., Salvucci and Entekhabi, 1994, for soil moisture or Scott et al., 1997, for individual components of evaporation) but may not encompass the entire range of possible interactions. Therefore, a feedback on precipitation through EF can, theoretically, result from any of the three components of $\lambda E$ (or a combination of them), all of them affected by antecedent precipitation itself.

As an extension to Fig. 3.1, Fig. 3.9 presents a schematic representation of the soil-moisture–precipitation feedback that distinguishes between the contributions of these three components of $\lambda E$. Precipitation impacts the three storage terms on different timescales, which might then impact EF and, thereby, precipitation, forming three interlinked feedback loops. The first loop ($C_1$–$A_1$–$B$) acts on a short timescale through $W_{\text{canopy}}$ and $E_{I}$, but is likely absent in our analysis due to the removal of days with morning rain. Indeed, field studies indicate high evaporation rates of intercepted water even at night (e.g., Pearce et al., 1980; Asdak et al., 1998; Holwerda et al., 2012), leading to complete evaporation of $W_{\text{canopy}}$ within a few hours. Therefore, evening precipitation is unlikely to provide intercepted water available during the before-noon time period (transparent black–green arrow in Fig. 3.9). Morning rainfall, on the other hand, may provide before-noon $W_{\text{canopy}}$ (black–green arrow) but given that such days may be of synoptic origin, they are excluded from our analysis, as noted previously. Moreover, in the few hours following rain, one may expect well-mixed conditions in the atmospheric boundary layer as well as low surface net radiation. Under these conditions, further rain would be more likely due to dynamical forcing from the atmosphere. The second loop ($C_2$–$A_2$–$B$) acts on a longer timescale, typically a few days, through $W_{\text{soil}}$ and $E_{\text{soil}}$. Finally, a third loop ($C_3$–$A_3$–$B$) acts on a mid- to long timescale, typically weeks to months, via $W_{\text{roots}}$ and $E_{\text{trans}}$. Ultimately, all three loops combine and act together on EF, which can impact precipitation. The distinction between these three components
3.6. SOIL MOISTURE AND INTERCEPTION EVAPORATION

Figure 3.9: Different water storage components contributing to $\lambda E$ and their potential relevance for afternoon convective precipitation. The letters ($A_i$, $B_i$, $C_i$) refer to the steps of the feedback loop shown in Fig. 3.1, where "$i$" indicates the evaporation component of concern (1 for evaporation from vegetation interception, $E_I$; 2 for bare soil evaporation, $E_{\text{soil}}$; 3 for plant transpiration, $E_{\text{trans}}$). The horizontal axis represents time, ending with day $i$, and precipitation over the past days to months is represented with its persistence timescales and its typical influence on the three water storage terms shown below: canopy or vegetation interception storage $W_{\text{canopy}}$ is affected by precipitation over the previous hours only ($C_1$). Surface soil moisture $W_{\text{top}}$ is impacted by precipitation in the previous days to weeks ($C_2$). Root-zone soil moisture $W_{\text{roots}}$ is mainly impacted by precipitation in the previous weeks to months ($C_3$). These three storages control their respective evaporation components, and thus $E_{\text{F}}$, in different regions. Over vegetated areas for interception ($A_1$), in a transitional soil-moisture–climate regime for soil evaporation ($A_2$), and in regions which are both vegetated and in a transitional climate regime for transpiration ($A_3$). Note that $A_2$ and $A_3$ can also occur in other regions in some circumstances (e.g., over wet regions, during dry years), and $W_{\text{roots}}$ includes $W_{\text{top}}$. Note that for loop 1 (through interception), a coupling cannot be distinguished from storm-scale precipitation persistence as before-noon interception is only expected in the presence of morning rain, mainly reflecting precipitation of synoptic origin. Precipitation over the previous evening usually does not affect before-noon $W_{\text{canopy}}$, but a transparent arrow is shown for rare cases where this might happen. Step B of the feedback remains a single component as the three evaporation components combine and only the total heat fluxes and their partitioning matter to precipitation occurrence.
has, to our knowledge, rarely been discussed in the literature in the context of EF–precipitation coupling or soil-moisture–precipitation feedback (with exceptions, e.g., Savenije, 1995b, 2004, for moisture recycling and Scott et al., 1995, 1997, for precipitation persistence). However, they may help to better understand some of our results.

In order to investigate the role of these three components, we compute \( \Delta \Gamma(X) \) using NARR data, where \( X \) is the water storage term controlling each component instead of EF. Storage terms are used instead of individual fluxes, since these are not available from NARR output. Figure 3.10a–d displays \( \Delta \Gamma \) in NARR computed with (a) EF, (b) surface soil moisture (for \( E_{\text{soil}} \)), (c) root-zone soil moisture (for \( E_{\text{trans}} \)), and (d) vegetation interception storage (for \( E_{I} \)). All these variables represent before-noon (9 a.m.–12 p.m.) values. The definition of surface and root-zone soil moisture in NARR is provided in Sec. 3.2.1.

Over the eastern US, most of the \( \Delta \Gamma \) signal found with EF does not appear to be related to soil moisture (neither for surface nor for root-zone soil moisture, except in Florida). This suggests that the EF variability is not driven by soil moisture variations in this region. On the other hand, \( \Delta \Gamma \) computed with morning vegetation interception storage displays a strong signal, suggesting that most of the signal with EF is linked to interception evaporation. However, \( \Delta \Gamma(\text{EF}) \) is not strongly sensitive to the exclusion of days with vegetation interception storage (Fig. 3.10e, while Fig. 3.10f displays the difference to the computation including all days and is rather small), despite the substantial fraction of days they represent in NARR (15–35%, Fig. 3.10g). Since the remaining signal (Fig. 3.10e) cannot be attributed to vegetation interception, it is likely either due to one of the remaining terms of evaporation or to atmospheric controls on EF through potential evaporation.

To test this hypothesis, the third row of Fig. 3.10 displays \( \Delta \Gamma(X) \) computed on days without vegetation interception and where \( X \) is (h) surface soil moisture, (i) root-zone soil moisture, and (j) potential EF (\( EF_{\text{pot}} = \lambda E_{\text{pot}}/(R_n - G) \), i.e., the EF that corresponds to potential evaporation from NARR, which is based on Penman–Monteith equation). Since results are noisy due to the low number of included days, Fig. B.5 in the Appendix displays the same analysis for the whole NARR time period (1979–2007). For most of the eastern US, \( \Delta \Gamma(\text{EF}_{\text{pot}}) \) appears to best reproduce the signal with EF on Supplement Fig. B.5, suggesting that atmospheric controls on EF (through \( \text{EF}_{\text{pot}} \)) at least partly induce the apparent positive coupling. Such a confounding effect could result from the control of temperature and humidity of the air mass on \( \text{EF}_{\text{pot}} \), which would then simply be a proxy for the likelihood of the air mass to produce rain, independently of surface fluxes.

Over other regions, we identify different key drivers based on Fig. 3.10 and Supplement Fig. chap4:suppl:f:suppl:narrstorages. Over the southwestern US, our analysis highlights surface and root-zone soil moisture as important
Figure 3.10: Identification of the drivers of the EF–precipitation relationship in NARR (see Fig. B.5 in the Appendix for the same analysis using the longer NARR time period). Top row: difference in the probability of afternoon rainfall, $\Delta \Gamma(X)$ on days with high vs. low $X$, where $X$ is the before-noon value of the different drivers. From left to right, $X$ is (a) EF and (b–d) the three water storage terms that control EF: (b) surface soil moisture ($W_{\text{top}}$, controls bare soil evaporation), (c) root-zone soil moisture ($W_{\text{roots}}$, controls plant transpiration), and (d) vegetation (canopy) interception storage ($W_{\text{canopy}}$, controls interception evaporation). Middle row: (e) $\Delta \Gamma(\text{EF})$ computation restricted to days without canopy storage, (f) difference between $\Delta \Gamma(\text{EF})$ computed with all days and with days without vegetation interception storage, and (g) percentage of days with interception storage. Bottom row: $\Delta \Gamma(X)$ restricted to days without interception storage, where $X$ is (h) surface soil moisture, (i) root-zone soil moisture, and (j) potential EF ($\text{EF}_{\text{pot}}$). High (low) $X$ refers to values higher (lower) than the 60th (40th) percentile of $X$, i.e., $\Delta \Gamma(X) = \Gamma(r|X > X_{Q60}) - \Gamma(r|X \leq X_{Q40})$. Values significantly different from 0 at the 90% level are indicated by a black asterisk. Grey dots indicate sites with no rainy days left.
Without interception (\( E_{\text{GLEAM}, P_{\text{NEXRAD}}} \))

Including interception (\( E_{\text{GLEAM}, P_{\text{NEXRAD}}} \))

\[
\begin{array}{cccc}
-0.15 & -0.05 & 0.05 & 0.15 \\
\end{array}
\]

\( TFS \)

significant (90%)

Figure 3.11: Influence of interception evaporation on TFS\(^*\) in the GLEAM–NEXRAD combination. Left: interception is not included in the EF computation. Right: interception is included in the EF computation and EF is then capped at 1. Values significantly different from 0 at the 90% level are indicated by a black asterisk. Empty dots indicate sites with unreliable NEXRAD data.

contributors, with interception playing a smaller role. Over the central US, no conclusion can be drawn from NARR as no EF–precipitation relationship is identified (see also Figs. 3.2 and 3.3).

As a sensitivity test, we also investigate the potential role of interception using GLEAM, where, in the default version, before-noon interception storage is neglected as it is based on a Gash analytical model (Gash, 1979), and therefore assumes that the vegetation water storage is evaporated within a model time step (Sec. 3.2.4). Here, we relax this assumption to allow comparison to results from NARR. Figure 3.11 displays TFS\(^*\) for the GLEAM–NEXRAD combination as shown earlier (standard version, left) and including interception evaporation from previous-day precipitation (right; see Sec. 3.2.4 for details on the computation). Including interception in that way leads to an increase in significant positive TFS\(^*\) signal, particularly over the eastern US, which is consistent with the results from NARR. However, we recall that this feature is likely not realistic: theoretical considerations do not support the presence of before-noon intercepted water storage in our analysis. Indeed, interception evaporation rates are high even at night (e.g., Pearce et al., 1980), as these are driven by advected energy or negative sensible heat flux rather than radiation (Pearce et al., 1980; Asdak et al., 1998; Holwerda et al., 2012). Thus, intercepted water evaporates within a few hours: for instance, a compilation of numerous studies on interception finds mean evaporation rates of 0.3 ± 0.1 mm h\(^{-1}\) and canopy storage of 1.2 ± 0.4 mm, leading to complete evaporation of the whole canopy reservoir in 4 ± 1.9 h (Miralles et al., 2010). The presence of interception storage during the before-noon time period is therefore largely restricted to days with morning precipitation, which are not included in our analysis as they are indicative of synoptic rainfall (Sec. 3.3.2).
Some exceptions might occur under very humid nighttime atmospheric conditions, preventing intercepted water from evaporating, but these cases likely coincide with morning precipitation and are thus more likely to represent the dynamical forcing from the atmosphere.

Thus, these theoretical considerations based on past field studies suggest an overestimation of the impact of interception in NARR, in line with the results from the validation of other climate reanalysis and land-surface models (e.g., Reichle et al., 2011; Van den Hoof et al., 2013; Davies-Barnard et al., 2014). This could be due to the parameterization of interception as a function of $E_{pot}$ (Chen et al., 1996), with interception unrealistically affected by net radiation (see Shuttleworth and Calder, 1979). Time series in NARR strongly support this hypothesis, showing that before-noon interception storage is most often provided by afternoon or evening precipitation on the previous day which does not evaporate in the night (see Appendix B.5 and Fig. B.6).

Overall, analyzing the role of individual components of $\lambda E$ in the relationship between EF and subsequent precipitation leads to similar conclusions in NARR and in GLEAM–NEXRAD. In the eastern US, vegetation interception evaporation and atmospheric controls on EF can lead to a likely overestimated relationship, due to, respectively, the fast rates of evaporation of intercepted water (see above) and the atmospheric origin of the signal. In the central and southwestern US, soil moisture (surface and root zone) drives the relationship where it exists, which would be consistent with the existence of a positive soil-moisture–precipitation feedback. These findings fit well with expectations based on climate regimes and vegetation cover: Fig. 3.12(a) highlights a wet regime in the eastern US, where land evaporation is controlled by radiation rather than soil moisture, unlike the soil-moisture–limited regime of the central and western US. In addition, vegetation interception is likely more relevant in the eastern US than in the central and southwestern US, as indicated by a high leaf area index in Fig. 3.12(b). Although we recall that the interception-related findings from NARR are not consistent with knowledge from field studies for the above-mentioned reasons, an impact of evening or nighttime interception evaporation via moisture recycling remains possible on longer timescales.

### 3.7 Discussion and conclusions

A recent study (Findell et al., 2011) statistically relates the occurrence of afternoon convective precipitation to before-noon evaporative fraction (EF) through the TFS metric (triggering feedback strength), based on data from the North American Regional Reanalysis (NARR), and suggests the existence of an extended region of positive land-surface–precipitation coupling over the eastern US. Our study extends that analysis with a systematic cross validation of additional, independent, observation-driven data sources and
Figure 3.12: (a) Land evaporation regime (blue for wet regime, red for transitional regime): multi-model analysis of controls on yearly land evaporation from Teuling et al. (2009a). Correlation between yearly evaporation and global radiation ($\rho_{R_g,E}$), and between yearly evaporation and precipitation ($\rho_{P,E}$), for the period 1986–1995. Each color corresponds to a unique combination of $\rho_{R_g,E}$ and $\rho_{P,E}$. (b) Mean summer (JJA) leaf area index $[m^2/m^2]$ over the period 1995–2007, from Stöckli et al. (2011).

Comparing the relationship patterns from the different data set combinations, the FLUXNET–NEXRAD and GLEAM–NEXRAD combinations do not reproduce positive TFS* in the eastern US found in NARR. Higher noise levels in these data sets and uneven sampling of different land cover types in the FLUXNET data may contribute to the differences. Nevertheless, our results suggest that land-surface dynamics in NARR and their stronger apparent coupling with precipitation in the eastern US might reflect model artifacts (see also Ferguson et al., 2012, who find that surface soil moisture from NARR correlates poorly with remote-sensing estimates in the eastern US). Conversely, a significant relationship between EF and convection triggering is found for the observation-driven GLEAM–NEXRAD combination in the central US (consistent with, e.g., Koster et al., 2004), although no such signal emerges from NARR in this region. The FLUXNET–NEXRAD combination displays low TFS* values there, possibly due to higher noise levels and short samples. Similarly, NARR might underestimate a possible EF–precipitation coupling in these regions. In the area of the southwestern US close to the Mexican border, all data sets agree on the existence of significant relationships between EF and convective triggering.

We find that the choice of the EF data set has a large impact on the relationship between EF and convection triggering, although the patterns of average EF, EF variability, and persistence in the different data sets do not clearly indicate the sources of this discrepancy. This comparison is further
hampered by short observational records, uncertainties, and different spatial scales.

Furthermore, we find that precipitation of the previous day is a better predictor of afternoon precipitation than before-noon EF, pointing to a short timescale dominance of the atmosphere over land. Although EF seems to provide a small additional predictability to precipitation alone, the confounding effects of precipitation on EF via soil moisture or intercepted water precludes definite conclusions on the existence of a land–precipitation coupling at this stage. Accounting for the individual components of land evaporation (plant transpiration, bare soil evaporation, and evaporation of intercepted water) in the analysis, we find that the coupling, if present, arises from distinct sources in different regions.

Over the eastern US, atmospheric controls on EF (i.e., the atmospheric demand through potential evaporation) and vegetation interception drive the EF–precipitation relationship in NARR. Atmospheric controls on EF might induce an apparent relationship, but identifying these drivers as, for example, in Aires et al. (2013) lies beyond the scope of this study. The unrealistic presence of before-noon intercepted water from previous evening rainfall in NARR, likely due to an underestimation of the rates of evaporation of intercepted water at night, may falsely contribute to the positive TFS over the eastern US, in line with the GLEAM–NEXRAD experiment (Fig. 3.11). This questions the reliability of NARR for these applications, despite its real advantage of high-quality precipitation assimilation. Other reanalysis products have issues with the representation of interception, e.g., MERRA (where the MERRA-Land product corrects for interception parameters among others; Reichle et al., 2011). These findings suggest a relatively short timescale of the EF–precipitation relationship in this region in NARR, which is consistent with the role of day-to-day precipitation persistence. Establishing a causal link between atmospheric- and interception-driven EF and precipitation is thus very difficult. Finally, we find that the EF–precipitation relationship found in NARR in the eastern US is not related to soil moisture, which makes sense given the humid climate regime with an expected low control of soil moisture on EF in this region, unlike what has been diagnosed in several studies for the central US (e.g., Koster et al., 2004; Teuling et al., 2009a; Seneviratne et al., 2010).

The processes in central and southwestern US are, indeed, different from those in the eastern US. Wherever significant positive relationships between EF and precipitation occurrence are found in GLEAM–NEXRAD or NARR, soil moisture is identified as the primary driver. This is consistent with the soil-moisture-limited evaporation regime in this transitional region (Koster et al., 2004; Seneviratne et al., 2010; Mueller and Seneviratne, 2012) and aligns well with expected regions of soil-moisture–precipitation coupling (e.g., Koster et al., 2004).
A number of processes are not considered in our analysis, such as the dominance of orographic lifting over land–atmosphere interactions over the northwestern US where evaporation is soil-moisture-limited (e.g., Schwalm et al., 2012), the effects from different land covers (e.g., young vs. mature forests; see Vickers et al., 2012), or other processes acting at smaller scales than those considered here. Detailed analysis of these local features is, however, beyond the scope and spatial scale of our study.

A small part of the signal in the eastern US in NARR cannot be attributed to vegetation interception, soil moisture, and EF\textsubscript{pot}. Nonlinear interactions between these variables possibly explain this signal, but the assimilation procedure may also affect it in a way that is difficult to assess. In GLEAM, the adjustments that were made to get estimates of before-noon EF introduce additional uncertainties that are difficult to quantify. Here we note that the reliability of the estimates is expected to decrease with increasing temporal resolution (Miralles et al., 2011b).

The presence of before-noon vegetation interception storage on days without morning precipitation is unlikely (based on previous field measurements, e.g., Pearce et al., 1980; Asdak et al., 1998; Holwerda et al., 2012). Thus, interception does not likely play a strong role for the triggering of convective storms via morning surface fluxes. Nonetheless, a direct impact via moisture recycling is possible and has already been suggested in the past (e.g., Savenije, 1995a,b). Additional moisture input to the atmosphere may thus provide more rainfall downwind on a longer timescale than the diurnal scale analyzed here. Indeed, evaporation from intercepted water has been estimated to amount to \(\sim 11\%\) of global land evaporation (Miralles et al., 2011a) and to 20–50\% over forests (e.g., Savenije, 2004; McLaren et al., 2008; Gerrits and Savenije, 2011).

The discrepancies between the coupling patterns of precipitation with soil moisture and EF, respectively, as well as the here-proposed explanations through interception evaporation and atmospheric controls on EF, have hardly been addressed in the recent literature on land–precipitation coupling (e.g., Findell and Eltahir, 2003a; Seneviratne et al., 2010; Findell et al., 2011; Taylor et al., 2011; Ferguson et al., 2012; Taylor et al., 2012). This adds to the complexity of this coupling but possibly explains some of the contradictions from recent studies (e.g., Findell et al., 2011; Ferguson et al., 2012; Taylor et al., 2012). We show that not only the individual segments of the soil-moisture–precipitation coupling (Fig. 3.1; Wei and Dirmeyer, 2010; Dirmeyer, 2011), but also the individual components of \(\lambda E\) may be crucial to uncover remaining uncertainties in land–atmosphere coupling.

Given the many unresolved issues in the investigation of land–precipitation coupling, further studies are required to pin down this complicated relationship. Analyses of the feedback accounting for precipitation persistence and confounding variables, applied to different temporal and spa-
tial scales and a wide range of data sets, are urgently needed. Moreover, improvements in models would allow for more realistic sensitivity studies. Finally, soil moisture and EF observations at scales relevant to land–atmosphere coupling (i.e., 10 km) would provide invaluable observational constraints on model results and understanding of land–atmosphere coupling.

Acknowledgments

We thank Han Dolman, Beverly E. Law, Altaf Alan, Markus Reichstein, Heini Wernli, and Christopher M. Taylor for discussions. We are also grateful for comments from the two anonymous reviewers. The authors acknowledge the Swiss State Secretariat for Education and Research SER for funding under the framework of the COST Action ES0804. NARR data is provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. This work used eddy-covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program – DE-FG02-04ER63917 and DE-FG02-04ER63911) and Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan). We acknowledge the financial support of the eddy-covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, University Laval, Environment Canada, and US Department of Energy, and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California – Berkeley, and the University of Virginia.

Edited by: L. Ganzeveld
Soil moisture-precipitation coupling: temporal and spatial perspectives from remote-sensing data

This chapter is an article in preparation
Benoit P. Guillod¹, Boris Orlowsky¹, Diego Miralles²,³, Adriaan J. Teuling⁴, and Sonia I. Seneviratne¹

Abstract Soil moisture is expected to influence precipitation via its controls on evaporation (i.e. moisture input to the atmosphere, direct impact) and the partitioning between the sensible and latent turbulent heat fluxes and resulting effects on the boundary layer stability (indirect impact). However, the sign of the soil moisture-precipitation coupling (and associated feedback) remains heavily debated, as the indirect impacts can theoretically lead to both signs of coupling. Recent studies have suggested the existence of negative feedback mechanisms in observations, contrasting with model results

¹Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland.
²Ghent University, Laboratory of Hydrology and Water Management, Ghent, Belgium.
³University of Bristol, School of Geographical Sciences, Bristol, UK.
⁴Wageningen University, Hydrology and Quantitative Water Management Group, Wageningen, the Netherlands.
and previous dominant findings from the literature. However, the proposed negative coupling relies on spatial analyses, which might investigate mechanisms related to spatial heterogeneity rather than a temporal feedback. Here we show that afternoon precipitation events tend to occur more often in wet conditions but are located over comparatively drier (less wet) patches. Using twelve years of remote-sensing data, we demonstrate that the use of spatial versus temporal methodologies leads to results of opposite signs in many regions. While we confirm the clear dominance of precipitation events occurring over soils that are drier than their surrounding, we highlight that in days when precipitation is found to occur, soils both at event locations and in the surroundings are wetter than average conditions in most regions. Although the apparent positive temporal coupling does not necessarily imply a causal relationship, these results potentially reconcile the notions of moisture recycling with local, spatially negative feedbacks. A positive temporal coupling might enhance precipitation persistence, while a negative spatial coupling tends to homogenize the land surface conditions.

4.1 Introduction

Land climate interactions play an important role in the climate system (Seneviratne et al., 2010), in particular in transitional climate regions where soil moisture can influence the partitioning of the energy available at the land surface into sensible ($H$) and latent ($\lambda E$) heat fluxes (Koster et al., 2004). Surface turbulent fluxes may influence precipitation directly via moisture input to the atmosphere (moisture recycling) as well as indirectly locally, via boundary layer dynamics and mesoscale circulations. Moisture recycling is expected to lead to a positive feedback, i.e. more precipitation induced by wet conditions. The indirect effect, on the other hand, can theoretically lead to feedbacks of both signs depending on atmospheric conditions (e.g. Findell and Eltahir, 2003a; Ek and Holtslag, 2004; Gentine et al., 2013).

Studies in the 1990s and 2000s have mostly identified positive coupling mechanisms using models or reanalyses (Schär et al., 1999; Koster et al., 2003; Guo et al., 2006; Findell et al., 2011). However, Taylor et al. (2012) have suggested a strong dominance of negative coupling mechanisms in observations contrasting with a strong positive coupling in Global Climate Models (GCMs). This negative coupling is likely driven by soil moisture-induced mesoscale circulations (Taylor et al., 2011). The apparent contradiction between these latter results and previous studies has led to a recent debate on positive versus negative feedback.

Since the sign of the feedback exhibited by climate models has been shown to be sensitive to the parameterization of convection (Hohenegger et al., 2009; Taylor et al., 2013), the use of models with explicit convection or – if possible – the direct inference of the underlying relationships from observations,
is essential to avoid parameterization-dependent results. Recently, global datasets of soil moisture, evaporation and precipitation from remote-sensing have become available and provide a unique opportunity to study the soil moisture-precipitation coupling mechanisms globally. However, observational analyses are impaired by the difficulty in establishing a causal relationship (e.g. Salvucci et al., 2002). The spatial analysis from Taylor et al. (2012) attempts to overcome this issue, and thereby demonstrates that precipitation occur more frequently over soils that are drier than the surrounding area.

Spatial analyses, however, might investigate processes that differ from the traditional understanding of soil moisture-precipitation feedback. Indeed, spatial gradients of soil moisture might be largely independent of large-scale moisture availability. Coupling mechanisms via mesoscale circulations (Taylor et al., 2011) might thus act on top of other effects such as moisture recycling, as suggested by Koster (2011). However, the large number of studies is difficult to compare due to the use of different datasets and methodologies for the investigation of soil moisture-precipitation coupling and feedbacks.

Here we use global remote-sensing based data and directly compare spatial and temporal approaches. We use the spatial method of Taylor et al. (2012) and a similar methodology for temporal relationships. We compare these two methodologies in order to assess whether the choice of an approach determines the sign of the feedback.

4.2 Data and Methods

4.2.1 Data

We use quasi-global (60°S-60°N) datasets of precipitation and soil moisture based on satellite remote-sensing observations over the period 2002-2011. As a primary precipitation dataset, we use estimates from CMORPH (Climate Prediction Center MORPHing method, Joyce et al., 2004) available at high spatial (0.25°x0.25°) and temporal (3h) resolution. CMORPH uses data from passive microwave sensor overpasses, which provide high-quality precipitation estimates but are available typically only several times a day, and it propagates these estimates based on motion-vectors derived from infra-red sensors onboard geostationary satellites, available at a high temporal resolution over most of the globe (see Appendix C.2.2 and Joyce et al., 2004, for further details). This dataset has been shown to agree well with rain gauges and other precipitation products (Joyce et al., 2004; Tian et al., 2009; Habib et al., 2012). Appendix C.3 provides results with other precipitation datasets (TRMM 3B42, Huffman et al., 2007, and PERSIANN, Hsu et al., 1997) as well as further details about these three precipitation datasets. Before conducting our analyses, the 3h precipitation data was adjusted to local time (LT, the local time based on longitude) by taking the closest 3h UTC-based time step.
GLEAM ("Global Land Evaporation: the Amsterdam Methodology", Miralles et al., 2011b) provides estimates of daily evaporative stress and evaporation components at a resolution of 0.25° based on remote-sensing data of radiation, precipitation, air temperature, soil moisture, vegetation optical depth and snow water equivalents. It has been extensively validated and inter-compared to other methodologies to estimate heat fluxes (e.g. Miralles et al., 2011a,b; Liu et al., 2013; Mueller et al., 2013; Trambauer et al., 2014). In particular, soil moisture from GLEAM has been successfully validated using measurements from 701 soil moisture sensors all across the world (see Supplementary Information of Miralles et al., 2014b). Here we use a modified version of GLEAM which provides estimates of morning (9:00LT) total evaporative stress $S$, defined through

$$E = SE_{\text{pot}}$$  \hspace{1cm} (4.1)

where $E$ denotes evaporation from the land surface and $E_{\text{pot}}$ is potential evaporation, estimated from the Priestley and Taylor approach (Priestley and Taylor, 1972) in GLEAM. $S$ therefore quantifies the land surface stress on evaporation from soil moisture and vegetation activity, and is defined as a linear function of soil moisture whose slope depends on vegetation optical depth. Further details about the GLEAM methodology are provided in the Appendix (C.2.1).

The GLEAM $S$ estimates present two major advantages over satellite-based surface soil moisture estimates: it includes the whole root zone (in addition to surface soil moisture), and its variability is limited to when soil moisture impacts evaporation. In the setup presented here, CMORPH is used as a precipitation input in GLEAM. Further details about the setup for GLEAM are provided in the Appendix (C.2.1), including a description of the other forcing datasets used. The sensitivity of our results to the chosen soil moisture dataset has been investigated by using satellite-based surface soil moisture from AMSR-E (Owe et al., 2008) as well as alternative GLEAM estimates based on different precipitation input (Appendix C.3). Note that we refer to $S$ as “soil moisture stress”, as the impacts of vegetation on $S$ are occurring over slower time scales than the one analyzed here.

### 4.2.2 Methods

To avoid the use of traditional methods that consist of analyzing time series separately at individual grid cells, we use the precipitation event detection technique from Taylor et al. (2012) and compare the pre-event soil moisture field of the events to a control sample based on non-event days. We define afternoon precipitation as the accumulated precipitation between 12-24LT, but results are highly similar for a choice of 12-21LT, used by Taylor et al. (2012) (Fig. C.3 in the Appendix). On a particular day, an event consists of 5x5 grid cells (0.25°x0.25° each, i.e. 1.25°x1.25° in total) centered at a
4.2. DATA AND METHODS

Location of local maxima (Lmax) where afternoon precipitation exceeds 4mm. The event domain is denoted Levt, while Lmin is the location of precipitation minimum within Levt.

A number of filters are applied to the individual 0.25° grid cells, and events for which Levt contains any filtered out grid cell are excluded from the computation. To ensure that events are generated in the analyzed afternoon, grid cells with morning (6-12LT) precipitation greater than 1mm are filtered out. Grid cells with fixed features that may influence the precipitation field (topography and water bodies) are also excluded (see Appendix C.1 and Taylor et al., 2012). In order to concentrate on the convective season, results in the main figures are presented for May-September latitudes North of 23°N, November-March latitudes South of 23°S, and including all months in the tropics where convection is the dominant process of precipitation generation (see e.g. Yang and Smith, 2008, for seasonal convective versus stratiform precipitation distributions). Results for individual seasons are available in the Appendix (Fig. C.11).

Once the events are detected, we analyze the corresponding patterns of morning soil moisture. Soil moisture anomalies $S'$ (i.e., with the mean seasonal cycle substracted) are used to mitigate the impacts of seasonality, and $S'_{Li}$ denotes $S'$ at a location $Li$. The metric to quantify the strength of the spatial relationship follows Taylor et al. (2012) and is applied for fixed 5°x5° boxes. For each event, we compute the difference in pre-event soil moisture $S'$ between Lmax and Lmin,

$$\Delta S'_e = S'_{Lmax} - S'_{Lmin}.$$  \hfill (4.2)

Then we define a control sample based on data for non-event days, $\Delta S'_c$ (using the same Lmax and Lmin locations on non-event days, and excluding days with morning rain). All values of $\Delta S'_e$ and $\Delta S'_c$ within a 5° box are pooled together, and we compute the difference in $\Delta S'$ values between the event and control sample,

$$\delta_e(\Delta S') = \text{mean}(\Delta S'_e) - \text{mean}(\Delta S'_c).$$  \hfill (4.3)

A null distribution of $\delta(\Delta S')$ is computed by pooling both samples ($\Delta S'_e$ and $\Delta S'_c$) together and taking 1000 bootstrap samples of a size equal to the size of $\Delta S'_e$. The quantile of the null distribution of $\delta(\Delta S')$ to which the actual $\delta_e(\Delta S')$ corresponds is a measure of the significance of the relationship.

To quantify the temporal relationship between morning soil moisture and afternoon precipitation, we adapt this methodology by using the soil moisture anomaly at Lmax (or Levt, Appendix C.4) instead of the difference $\Delta S'$, and thereby compute $\delta(S'_{Lmax}) = \text{mean}(S'_{Lmax,e}) - \text{mean}(S'_{Lmax,c})$ (or, similarly, $\delta(S'_{Levt})$), and the corresponding significance. This provides information about the soil moisture state on the morning of an event compared to the expectation. While this temporal approach is likely to be impacted
CHAPTER 4. GLOBAL LAND–PRECIPITATION COUPLING

**Figure 4.1:** Quantile of the spatial coupling metric \( \delta_c(\Delta S') = \text{mean}(\Delta S'_e) - \text{mean}(\Delta S'_c) \) under the Null hypothesis that no coupling exists, where \( \Delta S' = S'_{\text{Lmax}} - S'_{\text{Lmin}} \), the difference in \( S' \) between the location of rainfall maximum and the location of rainfall minimum. Low (high) quantiles indicate where \( \Delta S' \) is lower (higher) than expected. Horizontal black lines delineate the months included in the analysis: May-September North of 23°N (upper horizontal line), November-March South of 23°S (lower horizontal line) and all months in the tropics. Grey shading indicates non-significant relationships, grid cells with less than 25 events are left white. Corresponding results from Taylor et al. (2012) using surface soil moisture from AMSR-E and ASCAT and precipitation from CMORPH (version 0), as well as our results from various datasets, are shown in the Appendix (Fig. C.4).

by precipitation persistence at various time scales, it provides a temporal perspective that complements the spatial approach inherent to the metric of Taylor et al. (2012). Moreover, it avoids some of the issues with simple time-series analyses, where no distinction is made between advected and locally generated events.

### 4.3 Spatial and temporal results

The results of the spatial analysis are displayed in Fig. 4.1. Clearly, values of \( \delta_c(\Delta S') \) lie on the lower tail of the null distribution (low quantile values) at most locations, indicating that afternoon rain falls preferentially over soils that are drier than their surrounding. This is consistent with the findings from Taylor et al. (2012) (see also Fig. C.4a), while the slight differences in the regional patterns might be due to various factors (soil moisture data, precipitation dataset version, months used for the analysis, etc.). The dataset-related uncertainties are illustrated in the Appendix (Sec. C.3), where various dataset combinations are shown to lead to partly different patterns while all agreeing on the strong dominance of negative spatial coupling. These uncertainties have little impact to our interpretation as we precisely focus on the overall emerging signal.

The finding that afternoon rain falls preferentially over dry soils, how-
4.4 Discussion and conclusions

Our findings potentially reconcile a number of studies on soil moisture-precipitation feedback. Besides studies which consist of modeling experiments, most observation-based studies from the literature indeed investigate the coupling via temporal approaches, such as lagged correlations (e.g. between twice removed pentads, Koster et al., 2003, with a 1-day lag, Alfieri et al., 2008, or between daily variables at different times, Ferguson et al.,
2012) or metrics derived from lagged variables (e.g. morning evaporative fraction and afternoon precipitation in Findell et al., 2011). These temporal studies dominantly find positive relationships, i.e. more (frequent) rainfall occurring when soils are wet. Contrastingly, recent studies investigating the coupling by means of spatial approaches find negative relationships (Taylor et al., 2011, 2012), which might first appear contradicting. Here we show that this apparent contradiction is not related to the underlying data but to the chosen analysis approach.

We find that the soil moisture-precipitation coupling is apparently temporally positive but spatially negative. Note, however, that the apparent positive temporal coupling can be affected by precipitation persistence (e.g. Salvucci et al., 2002; Guillod et al., 2014), which could reflect persistence in large-scale controls such as atmospheric moisture advection. In such cases, it remains unclear whether soil moisture-precipitation feedback contributes to precipitation persistence and the observed positive temporal relationship, or if a weak or negative temporal coupling is hidden behind atmospheric persistence.

Spatial investigations of the coupling present the advantage of directly addressing persistence, as nearby locations are likely to exhibit similar conditions on the same day. However, such approaches by definition relate to spatial heterogeneity and might not fully address the question of whether a feedback occurs temporally (Koster, 2011), which might be most relevant for seasonal forecasting (e.g. Koster et al., 2010). Moreover, heterogeneity also exhibits some temporal variability: Section C.5 in the Appendix shows that precipitation events occur more often over heterogeneous conditions, and suggests that precipitation-generated heterogeneity possibly helps generating further precipitation events. This would lead to a positive feedback on a larger scale (Taylor et al., 2011). Note that a positive temporal coupling does not need to occur locally but could also affect areas downwind via moisture recycling (van der Ent and Savenije, 2011). Exhaustively addressing causality in temporal soil moisture-precipitation feedback has proven challenging over the past (Salvucci et al., 2002; Teuling et al., 2005) and will likely remain so until models are able to correctly capture the correct sign and strength of coupling metrics (Taylor et al., 2013).

In spite of the above-mentioned open questions, we demonstrate that the definition of soil moisture-precipitation coupling as a temporal or spatial property plays a crucial role in the obtained sign of relationship. Temporally, we find a dominance of apparent positive relationship, i.e. rain occurring more often in wet conditions, which might enhance precipitation persistence. Spatially, we find negative coupling, i.e. rain occurring more often over soils that are drier than the surrounding area, which might lead to an homogenization of water on land. If representative of causal relationships, these results would be consistent with the notion of moisture recycling (e.g. Eltahir and
Bras, 1996) as well as the existence of soil moisture-induced mesoscale circulations (Taylor et al., 2011). Improvements in models, in particular with respect to the representation of convection, is crucial to help disentangling the observed positive temporal relationship from atmospheric persistence. Although the mismatch in the sign of spatial coupling between Global Climate Models and observations (Taylor et al., 2012) does not necessarily imply a similar mismatch in temporal relationships, it still likely biases soil moisture fields and thereby possibly also affects temporal feedbacks.

Acknowledgments

We thank Christopher M. Taylor for helpful comments, as well as for providing data of a figure in his paper (Fig. C.4a in the Appendix) and the respective topography and water bodies masks (see Appendix C.1). We also thank Robert M. Parinussa for providing the daily AMSR-E data from LPRM in local time.
Conclusions and outlook

5.1 Conclusions

This thesis investigates two main aspects of land-climate interactions, that is, the relevance of the soil texture for the climate and the coupling between soil moisture and precipitation. For the first, we test the atmospheric response of an RCM to two soil maps with different geographical distributions of soil textures and, thereby, soil thermal and hydraulic properties. Second, we statistically investigate soil moisture-EF-precipitation coupling over North America and soil moisture-precipitation coupling globally in various observation-based datasets.

In Chap. 2, we could show that using soil maps from two different sources leads to substantial differences in the European summer climate as simulated by an RCM, locally reaching up to 2°C in mean 2m-temperature and 20% in precipitation. These differences are due to changes in EF resulting from different root-zone soil moisture stresses. They can be as large as typical differences between entirely different models from the PRUDENCE multimodel intercomparison. With additional simulations, we identify key soil parameters (field capacity, plant wilting point, hydraulic diffusivity) for these differences and highlight the importance of the vertical profile of soil moisture for $E$ and EF. Although the focus of this chapter lies on temperature, we also diagnose an overall positive feedback between soil moisture stress, EF, and precipitation.
In Chap. 3, analyses of temporal EF-precipitation coupling over North America highlight large uncertainties from the different datasets, in particular with respect to EF. For instance, remote-sensing based estimates indicate strong coupling in the Western US and no significant coupling over the Eastern US, while the NARR reanalysis displays strong coupling in the Eastern US. Short record lengths prevent clear conclusions from station-based data. We show that the temporal coupling is difficult to disentangle from precipitation persistence. However, analyzing the drivers of the EF-precipitation relationship provides interesting findings: in NARR, the coupling found over the Eastern US is driven by atmospheric controls on EF and vegetation interception rather than soil moisture. The interception-related relationship in NARR is further shown to be likely due to an underestimation of interception evaporation at night in this reanalysis due to its parameterization, providing unrealistic interception storage in the morning. In the remote-sensing estimates, similar unrealistic assumptions lead to the appearance of an EF-precipitation relationship in the Eastern US, while the Western US coupling is driven by soil moisture, which is consistent with the energy-limited evaporation regimes and high forest cover in the Eastern US versus the soil moisture-limited regime in the Western US.

The positive relationship over the US in the third chapter is consistent with the positive coupling from Chap. 2, although different regions are investigated and the imposed changes in soil texture in Chap. 2 are difficult to compare to the temporal coupling analysis of Chap. 3. In Chapter 2, most of the signal is driven by plant transpiration via root zone soil moisture. A small contribution from soil evaporation is seen as well. Interception is not represented in TERRA _ML, the land surface model of COSMO-CLM, due to numerical instabilities (note that these points are not discussed directly in Chap. 2). Thus, the feedback between EF and precipitation found in the RCM in Chap. 2 is likely occurring on longer time scales, via root zone soil moisture and transpiration, which is also expected from changing soil parameters. This may be comparable to the apparent coupling found in Chap. 3 with GLEAM over the Western US, which is suggested to also be caused by soil moisture, although the latter cannot be distinguished from precipitation persistence. Issues with interception identified over the Eastern US in NARR in Chap. 3, on the other hand, have no equivalent in the RCM study of Chap. 2 since interception is not modelled and, even if it were modelled, the experimental design would not directly impact it.

In our global analysis of the soil moisture-precipitation coupling (Chap. 4), we focus on the differences between temporal and spatial coupling between soil moisture and precipitation. The spatial coupling is found to be negative, consistently with Taylor et al. (2012), while the temporal coupling is globally apparently positive, but negative in some regions. Thus, although the temporal relationships can be impacted by persistence (e.g., Chap. 3),
we demonstrate that precipitation events tend to occur in wet conditions, but are located over drier (less wet) patches. We also show that precipitation events display a preference for heterogeneous soil moisture conditions. These results potentially reconcile various results from the literature, as temporal analyses generally display positive relationships, in contrast to negative coupling mechanisms identified from spatial studies. Although one cannot exclude that the temporal analyses are affected by atmospheric persistence, physical mechanisms would be consistent with the apparent contradiction between positive temporal and negative spatial coupling: Moisture recycling might provide moisture to the atmosphere, thereby leading to a positive temporal feedback, while mesoscale circulations induced by soil moisture spatial patterns lead to a spatially negative feedback and thereby to a homogenization of soil moisture spatially. Therefore, two parallel but different processes might in fact be investigated via the two types (temporal and spatial) of methodologies.

With the use of GLEAM data in both Chap. 3 and 4, the temporal coupling patterns from the two chapters could be expected to agree well over the US. However, this is not the case for two main reasons: First, the region over the Western US which exhibits positive relationships in Chap. 3 is masked by the topography filter in Chap. 4. Second, the negative temporal relationships found over the Central US in Chap. 4 is not detected in Chap. 3 due to the low availability of data (few sites, short records). The separation of different $E$ components and further confounding effects from Chap. 3 was not done in Chap. 4 as the focus of the latter is on the effect of soil moisture stress, and also because intercepted water is not part of GLEAM morning estimates and was shown to be a NARR feature in Chap. 3. Finally, the role of atmospheric persistence is not thoroughly investigated in Chap. 4, in spite of the potential strong role highlighted in Chap. 3. The main focus of Chap. 4 lies in the comparison of temporal and spatial coupling perspectives with an emphasis on the methodologies, with some mechanisms proposed but not always demonstrated.

A number of challenges remain. First, the scarcity of available observations is a long-standing issue. Therefore, the increasing availability of soil moisture and EF (or $E$) data, in particular from FLUXNET stations and satellite remote-sensing products, are most welcome and may allow for better constrained analyses. However, the so far large differences between different products suggest that these variables may remain poorly constrained for a while and that further progress is required.

Second, even with long records of high quality observations, confounding effects of precipitation persistence or other confounding variables often prevent a clear identification of the temporal (or, in some cases, spatial) coupling. Therefore, the systematic inclusion of potential confounding effects in coupling analyses, such as atmospheric conditions, large-scale forcing and
topography, is crucial. This is partly done in this thesis with respect to precipitation persistence, but it might require more thorough analyses. Within the many studies on this topic, only few actually account for persistence in a systematic way (e.g. Salvucci et al., 2002).

Finally, comparing, sorting and evaluating the many studies available in the literature is very difficult. This mostly relates to the investigation of the coupling via different metrics (TFS, $\delta_e$, lagged correlation, etc.) and variables (e.g. soil moisture, $E$ or EF), as well as datasets. To account for these uncertainty sources, the use of a wide range of datasets, variables and metrics is advisable in any further study on land-precipitation coupling. This would help to avoid potentially misleading dataset-dependent results (see Chap. 3 related to Findell et al., 2011) and allow for more direct comparability between studies.

5.2 Outlook

Here, I suggest possible future research topics, based on the results and conclusions of this thesis.

Role of soil organic carbon for climate and climate change: The sensitivity of the simulated climate to soil parameters analyzed in Chap. 2 relates to static maps of soil texture from different sources. A possible additional, more dynamic effect via soil organic carbon could also be investigated, as changes in soil organic carbon impact soil texture and parameters (Lawrence and Slater, 2008). Soil organic carbon is an important reservoir of carbon (Jobbágy and Jackson, 2000) which may exhibit large changes in the future (e.g. Friedlingstein et al., 2006). Current work in the context of a master thesis that I co-supervised (Hauser, 2013) suggests the existence of such impacts based on offline land surface model simulations, with processes relating among others to soil moisture profiles. These results also emphasize that the inclusion of dynamic soil organic carbon in GCMs may be useful to better account for all effects of soil organic carbon.

Land surface parameter calibration: Results from Chap. 2 highlight the importance of soil parameters and in particular hydraulic diffusivity for the simulated climate in COSMO-CLM. Soil parameters for a given soil class, as well as the geographical distribution of soil classes, are often only approximately defined and induce large uncertainties. This suggests that calibration of soil parameters may allow for bias reduction in RCMs and GCMs. For instance, relying on an objective calibration of regional climate models (Bellprat et al., 2012), recent work demonstrates that the soil hydrological conductivity/diffusivity is
5.2. OUTLOOK

a key parameter to constrain temperature biases in semi-arid regions over North America and Europe (Bellprat et al., 2013).

Climatic relevance of spatial and temporal coupling: If the temporal results from Chap. 4 of this thesis depict causal relationships, this chapter suggests that the processes underlying spatial and temporal coupling are different. Their respective climatic relevance remains to be determined. For instance, what is the actual role of temporal coupling in perpetuating extremes (e.g. droughts), and what is the relevance of spatial coupling for land surface heterogeneity? Ultimately, spatial and temporal coupling may interact in a way that determines subsequent precipitation events and their location. Understanding the extent to which these interactions impact current climate and climate change can potentially become very relevant. Future work based for instance on the methodology from Chap. 4 may clarify some of these points.

Contributions of remote control versus local coupling: Most studies investigate only one out of several soil moisture-precipitation physical mechanisms (e.g. moisture recycling or local coupling, the latter being the focus of Chap. 3 and 4). However, these likely interact strongly with one another. For instance, in cases with covariability between soil moisture at two locations \( L_1 \) and \( L_2 \), a local coupling signal detected at \( L_2 \) could in fact simply reflect moisture recycling via air advected from \( L_1 \). These different effects, shortly mentioned in Chap. 4, could be taken into account explicitly using moisture-tagging algorithms in combination with local coupling analyses. Distinguishing between locally generated rainfall and advected precipitation events potentially provides more sophisticated constraints for the removal of non-convective days in Chap. 3 and 4.

Cloud-resolving simulations: Given the numerous issues when diagnosing a coupling from observations (see e.g. Chap. 3 and 4), more realistic models may help to better understand the set of complex processes. Current model limitations relate mostly to the representation of convection (e.g. Guichard et al., 2004; Hohenegger et al., 2009), in particular in coarse-resolution GCMs and RCMs. The use of high-resolution, cloud-resolving simulations could allow to avoid this issue, with more accurate convection triggering and its relation with the atmospheric and surface variables and fluxes which are critical for land-precipitation interactions. At the same time, it may allow for the distinction between mesoscale circulation effects (spatial coupling) and the local triggering of convection (temporal coupling), which are not easy to disentangle from observations as shown in Chap. 4. While such investigation have already been undertaken for small domains (e.g. Schlemmer et al., 2012; Froidevaux
et al., 2014), their extension to larger regions is still constrained by CPU limitations. The future development of supercomputers may one day allow for global, long-term high resolution simulations, in which land-precipitation feedbacks could be investigated more realistically.

**Extreme precipitation events:** In the appendix of Chap. 4 (Appendix C.8), we find that, in general, soil moisture may locally impact the occurrence but not the amount of precipitation via local coupling. This finding, however, may not apply to extreme precipitation events with respect to moisture recycling. The quantification of the recycling contribution to extreme precipitation events may reveal some non-local impacts on precipitation amounts, e.g. through impacts on the total water content in the atmosphere and thus the potential amount of precipitation. This could potentially provide additional predictability with respect to the extent and amplitude of such extremes. Although we do not find any evidence for such an amplification of extreme precipitation in this thesis (maybe due to the chosen local coupling metric or to the limited quality of precipitation amount from remote-sensing products), analyses of moisture sources may well be able to capture such impacts.

**Validation of land-precipitation coupling in CMIP5 models:** Given the large uncertainty with respect to land surface processes in GCMs, an application of the land-precipitation coupling metrics to GCMs from the Coupled Model Intercomparison Project 5 (CMIP5) could be interesting to identify GCMs with unrealistic representations of these processes. Since spatial coupling has been shown to be poorly represented in these models (Taylor et al., 2012), analyses of temporal coupling may provide additional insights into the ability of models to represent land-precipitation feedbacks.
Appendix to Chapter 2

A.1 Parameterization of E and vertical water transport in TERRA_ML

This section described some selected aspects of TERRA_ML which are of particular relevance for our study. A more exhaustive documentation can be found at http://www.cosmo-model.org/content/model/documentation/core/default.htm.

A.1.1 Evapotranspiration

The parameterization of E is similar to that of the BATS model (Dickinson, 1984). Evapotranspiration includes the following components in TERRA_ML:

- Bare soil evaporation $E_b$
- Plant transpiration $E_p$
- Evaporation from interception and the snow reservoir

Interception evaporation is negligible, as well as evaporation from snow reservoir since we concentrate on summer in the analysis. We thus focus here on $E_b$ and $E_p$. 

109
Bare soil evaporation  \(E_b\) is parameterized as

\[
E_b = (1 - f_i) \cdot (1 - f_{\text{snow}}) \cdot (1 - f_{\text{plant}}) \cdot \text{Min}[-E_{\text{pot}}(T_{\text{sfc}}); F_m]
\]  

(A.1)

where \(F_m\) is the maximum moisture flux that the soil can sustain and \(f_{\text{plant}}\) is the fractional vegetation area and is given as an external parameter field. \(f_{\text{snow}}\) and \(f_i\) are the snow and interception fractional areas, respectively. \(F_m\) is parameterized as

\[
F_m = \rho_w C_k D \frac{s_t}{(z_u z_t)^{1/2}}
\]  

(A.2)

where \(z_u\), \(z_t\) and \(s_t\) are defined below (Eq. A.8 and surrounding text) and \(C_k\) is computed as

\[
C_k = 1 + 1550 \frac{D_{\text{min}}}{D_{\text{max}}} \cdot \frac{B - 3.7 + 5/B}{B + 5}
\]  

(A.3)

with \(B\) as defined for each soil class in Table 2.1 and

\[
D_{\text{min}} = 2.5 \cdot 10^{-10} \text{m/s}^2
\]  

(A.4)

\[
D_{\text{max}} = B \Phi_0 K_0 / \rho_{wm}.
\]  

(A.5)

Here \(\Phi_0 = 0.2\ m\) is the soil water suction at saturation and \(\rho_{wm} = 0.8\) is the fraction of saturated soil filled by water, while \(B\) and \(K_0\) depend on the soil class (see Table 2.1). \(D\) is expressed as

\[
D = 1.02 D_{\text{max}} s_u^{B_f^2} (s_t / s_u)^{B_f}
\]  

(A.6)

with \(B_f\) given by

\[
B_f = 5.5 - 0.8 B \left[1 + 0.1 (B - 4) \log_{10} \frac{K_0}{K_R}\right]
\]  

(A.7)

with \(K_R = 10^{-5}\text{m/s}\).

In equation A.2 and A.6, \(s_u\) and \(s_t\) are average values of soil water content normalized by the volume of void (\(\theta_{PV}\)) for two layers corresponding to depths of \(z_u = 0.09\ m\) and \(z_t = 1\ m\). These layers approximate Dickinson’s layers (0 – 0.1\ m and 0 – 1\ m) by setting the lower boundary (layers number \(n_u\) and \(n_t\) counting from the surface, respectively) as the lowest layer for which the lower boundary does not exceed 0.1\ m and 1\ m, respectively.

\[
s_{\text{u},t} = \frac{\sum_{k=1}^{n_{\text{u},t}} W_k}{\theta_{PV} \sum_{k=1}^{n_{\text{u},t}} \Delta z_k}
\]  

(A.8)

where \(s_{\text{u},t}\) and \(n_{\text{u},t}\) denote either \(s_u\) and \(n_u\) or \(s_t\) and \(n_t\), and \(W_k\) is the water content of layer \(k\) (in meters).
Plant transpiration $E_p$ is parameterized as

$$E_p = f_{\text{plant}} \cdot (1 - f_i) \cdot (1 - f_{\text{snow}}) \cdot E_{\text{pot}}(T_{\text{sfc}}) r_a (r_a + r_f)^{-1}$$

(A.9)

i.e. similarly to Dickinson (1984) but with additional assumptions (see online documentation for more details). Here, atmospheric resistance $r_a$ is given by

$$r_a^{-1} = C_d |v_h| = C_A$$

and foliage resistance is given by

$$r_f^{-1} = r'_C = C'_A$$

with $C_F = f_{\text{LAI}} r_{\text{la}}^{-1}$, $r_{\text{la}}^{-1} = C' u_{\text{a}}^{1/2}$ and $r' = r_{\text{la}} (r_{\text{la}} + r_s)^{-1}$. $f_{\text{LAI}}$ is the leaf area index, given as an external parameter, and the stomatal resistance $r_s$ is defined by

$$r_{s}^{-1} = r_{\text{max}}^{-1} + (r_{\text{min}}^{-1} - r_{\text{max}}^{-1})[F_{\text{rad}} F_{\text{wat}} F_{\text{tem}} F_{\text{hum}}]$$

(A.10)

with $r_{\text{min}} = 150 s/m$ and $r_{\text{max}} = 4000 s/m$. The functions $F$ describe the influence of the following conditions on the stomatal resistance: radiation ($F_{\text{rad}}$), soil water content ($F_{\text{wat}}$), ambient temperature ($F_{\text{tem}}$) and ambient specific humidity ($F_{\text{hum}}$), with $F = 1$ for optimal conditions and $F = 0$ for unfavorable conditions. In particular, we note the function describing the water limitation:

$$F_{\text{wat}} = \text{Max} \left[ 0; \text{Min} \left( 1; \frac{\theta_{\text{root}} - \theta_{\text{PWP}}}{\theta_{\text{TLP}} - \theta_{\text{PWP}}} \right) \right]$$

(A.11)

where $\theta_{\text{root}}$ is the liquid water content fraction of the soil averaged over the root depth, $\theta_{\text{PWP}}$ is the permanent wilting point (see Table 2.1) and $\theta_{\text{TLP}}$ is the turgor loss point of plants, parameterized following Denmead and Shaw (1962) as

$$\theta_{\text{TLP}} = \theta_{\text{PWP}} + (\theta_{\text{FC}} - \theta_{\text{PWP}}) \cdot (0.81 + 0.121 \arctan(E_{\text{pot}}(T_{\text{sfc}}) - E_{\text{pot,norm}})))$$

(A.12)

with $E_{\text{pot,norm}} = 4.75 \text{mm/d}$. For these two components, potential evaporation $E_{\text{pot}}$ is expressed as:

$$E_{\text{pot}}(T_{\text{sfc}}) = \rho C_d |v_h| (q^v - Q^v(T_{\text{sfc}}))$$

(A.13)

where $T_{\text{sfc}}$ is the temperature at the surface (uppermost soil layer for both $E_b$ and $E_p$), and $Q_v$ is the saturation specific humidity. $|v_h|$ is the absolute wind speed the the lowest grid level above the surface and $C_d^v$ is the bulk-aerodynamical coefficient for turbulent moisture transfer, calculated diagnostically.

### A.1.2 Vertical soil water transport

The vertical water transport is based on Richards equation, which in the vertical direction is usually expressed as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[ K_w \left( \frac{\partial \psi}{\partial z} + 1 \right) \right]$$

(A.14)
where $\theta$ is the soil water content and $\psi$ is the water potential. On the right side of Equation A.14, $\frac{\partial \psi}{\partial z}$ refers to capillary forces, while 1 represents gravity. However, in TERRA_ML, this equation is expressed using only $\theta$ and not $\psi$. To do so, hydraulic diffusivity is introduced as $D_w = K_w \frac{\partial \theta}{\partial z}$ and thus

$$K_w \frac{\partial \psi}{\partial z} = K_w \frac{\partial \psi}{\partial \theta} \frac{\partial \theta}{\partial z} = D_w \frac{\partial \theta}{\partial z} \quad \text{(A.15)}$$

This leads to the equations used in TERRA_ML, where the soil water flux is expressed as

$$F = -\rho_w \left[ -D_w \frac{\partial \theta}{\partial z} + K_w \right] \quad \text{(A.16)}$$

and the change over time in soil water content in each layer is defined as

$$\frac{\partial \theta}{\partial t} = \frac{1}{\rho_w \frac{\partial F}{\partial z}} \quad \text{(A.17)}$$

Here, the vertical transport due to gravity and capillary forces is represented by $K_w$ and $D_w$, respectively. Note that hydraulic conductivity and hydraulic diffusivity represent the same physical characteristics of the soil, namely the ability of water to flow into it, but they express it in different units. The presence of both variables is specific to this modelling approach. In some other land-surface models (e.g. Community Land Model, see Lawrence et al., 2011), water potential is used and hydraulic diffusivity does not appear.

In addition, runoff is parameterized for any layer with $\theta > \theta_{FC}$ and a negative divergence of the fluxes (A.16).

Hydraulic diffusivity $D_w$ and hydraulic conductivity $K_w$ depend on the water content $\theta$ as:

$$D_w(\theta_l) = D_0 \exp \left( D_1 \frac{(\theta_{PV} - \theta_l)}{(\theta_{PV} - \theta_{ADP})} \right) \quad \text{(A.18)}$$

$$K_w(\theta_l) = K_0 \exp \left( K_1 \frac{(\theta_{PV} - \theta_l)}{(\theta_{PV} - \theta_{ADP})} \right) \quad \text{(A.19)}$$

$\theta$ is defined for each layer as $\theta_l = \frac{W_l}{\Delta z_l}$.

### A.2 Conversion of the soil maps

#### A.2.1 FAO and JRC soil maps: raw data

**FAO** The Soil map of the World released by the FAO is available at a resolution of 5 arc minutes and in geographical projection. The raw data used is taken from the Digital Soil Map of the World cd-rom (see [http://www.fao.org/icatalog/search/dett.asp?aries_id=103540](http://www.fao.org/icatalog/search/dett.asp?aries_id=103540)). The classification considers three classes reflecting soil texture: coarse, medium, and fine. For use in TERRA_ML, data from the top layer of the soil is considered.
A.2. CONVERSION OF THE SOIL MAPS

<table>
<thead>
<tr>
<th>JRC soil class (TEXT_SRF_DOM)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>No information</td>
<td>-</td>
</tr>
<tr>
<td>Coarse</td>
<td>18% &lt; clay and &gt; 65% sand</td>
</tr>
<tr>
<td>Medium</td>
<td>(18% &lt; clay &lt; 35% and &gt;= 15% sand) or (18% &lt; clay and 15% &lt; sand &lt; 65%)</td>
</tr>
<tr>
<td>Medium fine</td>
<td>&lt; 35% clay and &lt; 15% sand</td>
</tr>
<tr>
<td>Fine</td>
<td>35% &lt; clay &lt; 60%</td>
</tr>
<tr>
<td>Very fine</td>
<td>clay &gt; 60%</td>
</tr>
<tr>
<td>No mineral texture</td>
<td>Peat soils</td>
</tr>
</tbody>
</table>

Table A.1: Categories of the attribute “TEXT_SRF_DOM” in the JRC soil map.

JRC The Soil Geographical Database of Eurasia (SGDBE, see Lambert et al., 2002) at a scale of 1:1,000,000 is a digitized European map of the soil and related attributes. It is part of the European Soil Database, a product released in 2006 by the JRC (Morvan et al., 2008; Panagos et al., 2012) and available at [http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/index.htm](http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/index.htm). It contains a large number of attributes, of which two are used in this study. These two attributes reflect the properties of the top layer of the soil, thus being consistent with the data used from the FAO soil map. The soil class is derived from the attribute “dominant surface textural class of the STU” (“TEXT_SRF_DOM”). This attribute contains the classes listed in Table A.1. Non-soil classes are derived from the attribute “Soil major group code of the STU from the 1990 FAO-UNESCO Soil Legend” (“FAO90-LEV1”), which contains 28 soil categories and 6 non-soil categories. Non-soil categories are listed in Table A.2. More details about the attributes is given at [http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/popup/sg_attr.htm](http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/popup/sg_attr.htm).

A.2.2 Conversion to TERRA_ML format: resolution and classes

FAO The conversion of the raw data into TERRA_ML classes with the desired resolution is done using the PrEProcessor of time invariant parameters (PEP) of COSMO-CLM (Smiatek et al., 2008). In this code, the number of points for each textural class (coarse, medium, fine) within a grid cell is determined and a weighted mean texture is computed. The assigned class in TERRA_ML is a function of this weighted mean. Ice, rock and peat are
Table A.2: Non-soil categories of the attribute “FAO90-LEV1” in the JRC soil map and their conversion into TERRA_ML classes. Note that this attributes contains 28 soil categories as well, such as Acrisol, Alisol, etc. These are not considered in the present study.

<table>
<thead>
<tr>
<th>JRC non-soil class</th>
<th>TERRA_ML class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Town</td>
<td>Rock</td>
</tr>
<tr>
<td>Soil disturbed by man</td>
<td>Rock</td>
</tr>
<tr>
<td>Water body</td>
<td>Rock</td>
</tr>
<tr>
<td>Marsh</td>
<td>Rock</td>
</tr>
<tr>
<td>Glacier</td>
<td>Ice</td>
</tr>
<tr>
<td>Rock outcrops</td>
<td>Rock</td>
</tr>
</tbody>
</table>

For input into COSMO-CLM, the JRC data was first resampled from its original 1km resolution in Lamberts azimuthal projection to 1 arc second resolution in geographical projection by nearest neighbor interpolation. Classes were then converted to corresponding classes in TERRA_ML. For non-soil classes, the category “glacier” was converted to the class “ice” in TERRA_ML, while all other non-soil classes were converted to “rock”, as shown in Table A.2. However, non-soil classes were attributed only where there was no data available about the soil class in the attribute “TEXT_SRF_DOM” (i.e. where the value is 0). For soil classes, a conversion scheme was defined by refering to the look-up table described by Smiatek et al. (2008) and comparing it to the definition of the classes in Table A.1 based on the proportion of clay, sand and silt. The resulting conversion scheme is shown in Table A.3. Comparing the legend of the two classes in Table A.3 gives us confidence in the chosen scheme. Note, however, that there is no ideal conversion scheme. For instance, here no class in the JRC data is converted to the TERRA_ML class “sandy loam” (coarse to medium). Conversely, the two classes “fine” and “very fine” in JRC are converted to the same class in TERRA_ML (clay, i.e. fine). These two cases illustrate the difficulty to translate a soil dataset with given soil categories into other categories, and, therefore, the associated uncertainties. Grid points where no information on both soil and non-soil categories was available (e.g. over
A.2. CONVERSION OF THE SOIL MAPS

### Table A.3: Conversion table of the soil classes from JRC (attribute “TEXT_SRF_DOM”) to TERRA_ML. Note that the scheme applied does not lead to any grid point with the soil class “sandy loam” in TERRA_ML. The columns for TERRA_ML follows Doms et al. (2011).

<table>
<thead>
<tr>
<th>JRC soil class (TEXT_SRF_DOM)</th>
<th>TERRA_ML class</th>
<th>short name (soil class)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarse</td>
<td>Coarse sand</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>Medium loam</td>
<td></td>
</tr>
<tr>
<td>Medium fine</td>
<td>Medium to fine loamy clay</td>
<td></td>
</tr>
<tr>
<td>Fine</td>
<td>Fine clay</td>
<td></td>
</tr>
<tr>
<td>Very fine</td>
<td>Fine clay</td>
<td></td>
</tr>
<tr>
<td>No mineral texture</td>
<td>Histosols peat</td>
<td></td>
</tr>
</tbody>
</table>

North Africa) were filled using the original FAO soil map.

Finally, the aggregation at the model resolution (0.44° in our case) is done using the COSMO-CLM PEP program as described by Smiatek et al. (2008). The method used in this program is the majority approach, i.e. the soil class with the higher number of points within a grid cell is attributed to that grid cell.
Appendix to Chapter 3

B.1 North American Regional Reanalysis: sensitivity of EF-precipitation relationship to the time period and to the selection of potentially convective days

Although the NARR product is available from 1979 to present, we use only a sub-period with available data in other datasets. Since most of the analyses presented in the main part of this paper include data from both NARR and the GLEAM-NEXRAD combination, computations are restricted to days when data is available from the GLEAM–NEXRAD combination (years 1995–2007 minus some gaps). In order to test the impact of the reduced amount of data, Fig. B.1 displays $TFS^*$ computed from NARR using years 1979–2007 (left) and using only years and days with data in the GLEAM–NEXRAD combination (right). This leads to a reduced number of sites with significant positive $TFS^*$ and to slightly lower values. Nonetheless, the overall patterns remain highly similar.

In addition, the impact of the criteria for the selection of convective days (see Sec. 3.3.2) is tested in Fig. B.2, showing results for our radiation-based criterion (left) and (right) the original criterion from Findell et al. (2011). The patterns of $TFS^*$ are highly similar using either criterion, albeit our criterion yields slightly lower values. The radiation criterion is used in this study in order to allow for the same criterion for all datasets.
Figure B.1: Impact of the time period on the computed impact of EF on convection triggering in NARR: TFS\(^\star\) computed with (left) data from 1979–2007 and (right) with the subset of days when data is also available from the GLEAM–NEXRAD combination (years 1995–2007 minus some periods with gaps). Similar results are found for both subsets of data, highlighting a small impact of the time period on the results.

Figure B.2: Impact of the selection of potentially convective days on the computed impact of EF on convection triggering in NARR: TFS\(^\star\) computed using (left) our radiation-based criterion and (right) CTP-based criterion from Findell et al. (2011). Days with morning precipitation are removed in both cases. The computation is restricted to 1979–2003 due to readily available CTP data over that period. The resulting maps show that the use of our radiation-based criterion instead of the original CTP-criteria leads to highly similar values of TFS\(^\star\).
B.2. GLEAM: SENSITIVITY TO INPUT DATA

As described in the main part of this paper (see Section 3.2.4), GLEAM-derived estimates of before-noon EF are computed using precipitation from NEXRAD as input. Other precipitation products could be used, but have to match the required timing described in Sec. 3.2.4. Table B.1 shows the definition of days for two common daily precipitation datasets, CPC-Unified (Chen et al., 2008) and GPCP-1DD (Huffman et al., 2001). While GPCP-1DD has a clearly defined timing that is consistent globally, it does not fit our experimental setting with GLEAM, since precipitation during roughly the 12 hours preceding the estimated EF would be missed. CPC-Unified, on the other hand, fits our requirements over the US (day ending in the morning), although one cannot exclude that afternoon precipitation is included due to possibly different reporting times between different PIs.

We restricted the analyses in the main part of the paper to a version of GLEAM forced by NEXRAD precipitation, but Fig. B.3 shows results with CPC-Unified and PERSIANN (a 3-hourly satellite product, Hsu et al., 1997), two other suitable precipitation forcings. For the TFS* computation, NEXRAD precipitation is used in all cases. All three estimates display similar West-East gradients, highlighting the robustness of our GLEAM-based results with respect to precipitation input. Nonetheless, TFS* values tend to be lower and less significant with CPC and PERSIANN.

Figure B.3: Triggering Feedback Strength (TFS*) based on remote-sensing products (GLEAM, NEXRAD) where EF is computed using different precipitation datasets as input (NEXRAD, CPC, PERSIANN). Note that precipitation from NEXRAD is used for TFS computation for all three datasets.

B.2 Sensitivity of GLEAM to the precipitation dataset used as input

As described in the main part of this paper (see Section 3.2.4), GLEAM-derived estimates of before-noon EF are computed using precipitation from NEXRAD as input. Other precipitation products could be used, but have to match the required timing described in Sec. 3.2.4. Table B.1 shows the definition of days for two common daily precipitation datasets, CPC-Unified (Chen et al., 2008) and GPCP-1DD (Huffman et al., 2001). While GPCP-1DD has a clearly defined timing that is consistent globally, it does not fit our experimental setting with GLEAM, since precipitation during roughly the 12 hours preceding the estimated EF would be missed. CPC-Unified, on the other hand, fits our requirements over the US (day ending in the morning), although one cannot exclude that afternoon precipitation is included due to possibly different reporting times between different PIs.

We restricted the analyses in the main part of the paper to a version of GLEAM forced by NEXRAD precipitation, but Fig. B.3 shows results with CPC-Unified and PERSIANN (a 3-hourly satellite product, Hsu et al., 1997), two other suitable precipitation forcings. For the TFS* computation, NEXRAD precipitation is used in all cases. All three estimates display similar West-East gradients, highlighting the robustness of our GLEAM-based results with respect to precipitation input. Nonetheless, TFS* values tend to be lower and less significant with CPC and PERSIANN.
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Timing definition (UTC) for day $i$</th>
<th>US local timing West (East) coast</th>
<th>Reference</th>
<th>Shown in</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC-Unified</td>
<td>Country-dependent</td>
<td>4AM (7AM), day $(i-1)$ to 4AM (7AM), day $(i)$</td>
<td>Chen et al. (2008)</td>
<td>Fig. B.3</td>
</tr>
<tr>
<td></td>
<td>USA: from 12z, day $(i-1)$ to 12z, day $(i)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPCP-1DD</td>
<td>UTC day (0z-0z)</td>
<td>4PM (7PM), day $(i-1)$ to 4PM (7PM), day $(i)$</td>
<td>Huffman et al. (2001)</td>
<td>not suitable</td>
</tr>
<tr>
<td>PERSIANN</td>
<td>3-hourly resolution</td>
<td></td>
<td>Hsu et al. (1997)</td>
<td>Fig. B.3</td>
</tr>
<tr>
<td>NEXRAD</td>
<td>3-hourly resolution</td>
<td></td>
<td><a href="http://www.ncdc.noaa.gov/oa/radar/radarresources.html">http://www.ncdc.noaa.gov/oa/radar/radarresources.html</a></td>
<td>main paper</td>
</tr>
</tbody>
</table>

**Table B.1:** Definition of the day (timing) for the discussed precipitation datasets. PERSIANN and NEXRAD are available at a 3-hourly resolution and are used in the analysis. Two common daily precipitation datasets, CPC-Unified and GPCP-1DD are described as follows: For precipitation on day $(i)$, the timing definition in UTC and US local time are shown. The US local timings for the West coast and the East coast are expressed in standard time (i.e., local time without the offset for daylight saving time).
B.3 EF Correlations at longer time scales

Figure 3.5 in Chap. 3 highlights weak correlations between before-noon EF estimates from different datasets. Figure B.4 displays correlations of 10-days and monthly means of before-noon EF values. Correlations are higher than for daily values of before-noon EF. Nonetheless, correlations remain small, mostly around 0.5, highlighting uncertainties in EF even on longer time scales. The correlations become less significant for longer time scales due to the lower number of data points. Using EF anomalies instead of absolute values leads to similar results, albeit with slightly lower correlations (not shown).

![Figure B.4: Correlations of JJA before-noon EF between different datasets for (a) 10-days averages and (b) monthly averages. The size of the dots indicates the number of days included in the computation, and significant correlations at a 99% level are indicated by a black star. Empty dots for GLEAM indicate sites with unreliable NEXRAD (and thus GLEAM) data.](image-url)
B.4 Decomposition of the NARR signal: full NARR time period

Section 3.6 decomposes the EF-precipitation relationship from NARR into its components of land evaporation and the atmospheric forcing (Fig. 3.10). While in the paper all analyses with NARR are conducted for days also available in GLEAM-NEXRAD for consistency, the analysis of Fig. 3.10 can easily be extended to include the whole NARR time period (1979–2007). This is particularly useful for the last row of Fig. 3.10, where the low number of days without interception storage leads to noisy results. Fig. B.5 shows the results of the computations for the 1979-2007 period and confirms that, once the effect of interception storage is removed, the atmospheric forcing on EF \( EF_{pot} \) seems to play a greater role than soil moisture over the Eastern US.
B.4. DECOMPOSITION OF THE NARR SIGNAL

Figure B.5: Same as Fig. 3.10 in Chap. 3 but for the longer NARR time period (1979–2007): Identification of the drivers of the EF-precipitation relationship in NARR. Top row: difference in the probability of afternoon rainfall, $\Delta \Gamma(X)$ on days with high vs. low $X$, where $X$ is the beforenoon value of the different drivers. From left to right, $X$ is (a) EF and (b–d) the three water storage terms that control EF: (b) surface soil moisture ($W_{top}$, controls bare soil evaporation), (c) root zone soil moisture ($W_{roots}$, controls plant transpiration) and (d) vegetation (canopy) interception storage ($W_{canopy}$, controls interception evaporation). Middle row: (e) $\Delta \Gamma(EF)$ computation restricted to days without canopy storage, (f) difference between $\Delta \Gamma(EF)$ computed with all days and with days without vegetation interception storage, and (g) percentage of days with interception storage. Bottom row: $\Delta \Gamma(X)$ restricted to days without interception storage where $X$ is (h) surface soil moisture, (i) root zone soil moisture and (j) potential EF ($EF_{pot}$). High (low) $X$ refer to values higher (lower) than the 60th (40th) percentile of $X$, i.e. $\Delta \Gamma(X) = \Gamma(r|X > X_{Q60}) - \Gamma(r|X \leq X_{Q40})$. Values significantly different from 0 at the 90% level are indicated by a black star. Grey dots indicate sites without days with afternoon precipitation.
B.5 Interception in NARR: before-noon storage source

GLEAM assumes that interception storage at the time of the estimated EF (9-12LT) is negligible, because days with morning (6-12LT) precipitation are removed from the analysis and intercepted water is known to evaporates within a few hours, even at night. However, our analysis from Section 3.6 (see also Sec. B.4 in this appendix) suggests that there is some intercepted water left during the before-noon time period in a substantial number of days in NARR, and that this could explain part of the EF-precipitation relationship over the Eastern US. In Sec. 3.6, we mention that this interception storage might be provided by precipitation on the previous day, a possibly non-realistic feature. Indeed, like many models, the Noah land surface model used in NARR uses a parameterization of interception evaporation which relies on potential evaporation (i.e., mostly net radiation), leading to low (or no) evaporation at night. This would be inconsistent with field studies, which have shown that interception evaporation is typically as high during the night as during the day and suggest that advected energy or negative sensible heat flux, rather than radiation, drive interception evaporation (e.g., Pearce et al., 1980; Asdak et al., 1998; Holwerda et al., 2012).

To support our hypothesis, Fig. B.6 displays composites of time series of the interception storage (left) and precipitation (right) in NARR, for 6 representative sites in the Eastern US and for days without morning precipitation. As expected, the interception storage is filled up by rainfall during the afternoon or evening of the previous day. In spite of low precipitation during the night (in most cases, no precipitation at all), intercepted water does not evaporate until the before-noon time period, supporting our hypothesis. Note that the interception storage capacity in NARR is only 0.5 mm, and that this amount would re-evaporate within a few hours (e.g. in less than 1.5 h with the night-time interception evaporation rate of 0.37 mm/h from Pearce et al., 1980).

Thus, the parameterization of interception in NARR appears to underestimate the evaporation of intercepted water at night, which leads to an overestimation of morning interception storage and related evaporation.
Figure B.6: Composite time series of interception storage (left) and precipitation (right) in NARR at 6 sites in the Eastern US, on days without morning precipitation (< 1 mm from 6-12LT). The before-noon time period is hatched in grey and the previous 24 h are displayed (day and hour of the day indicated on the horizontal axes), with midnight indicated by a vertical black line. The thick orange/blue lines indicate the median of all cases, while color shading indicates percentiles of the time series according to the legend shown at the top. Interception storage capacity (i.e., maximum value) is defined as 0.5 mm in NARR. In most cases, interception storage is filled in the afternoon or evening of the previous day and does not re-evaporate at night in spite of no or low precipitation.
Appendix to Chapter 4

C.1 Precipitation events detection technique

Similarly to Taylor et al. (2012), precipitation events are defined as 5x5 grid cells (0.25°x0.25° each, i.e. 1.25°x1.25° in total, referred to as Levt) centered at a location of local precipitation maxima (Lmax) where afternoon precipitation (12-24LT) exceeds 4mm. In case of overlap between several events on the same day, only the event with the largest accumulated precipitation at Lmax is retained. Morning precipitation is not allowed to exceed 1mm in any of the 25 grid cells within Levt. Furthermore, fixed features that may influence the precipitation field are removed: grid cells with a range of topographic height within a box of 1.25° exceeding 300m, as well as grid cells where water bodies cover over 5% of the area are removed as in Taylor et al. (2012), using the datasets therein. The locations of precipitation minimum within Levt are denoted Lmin, and in case of multiple grid cells belonging to Lmin for a given event the soil moisture value at Lmin is defined as the average of all the corresponding values.

Figure C.1 illustrates the event definition with an example day over West Africa. Events are delimited in black, with their respective Lmax indicated by letters and Lmin indicated by grey dots. Black symbols indicate filters from morning precipitation (crosses), topography (triangle) and water bodies (circles). The two events delimited in grey are not included in the computation as they include grid cells filtered out due to topography (event “A”) or
Figure C.1: Example events in West Africa, on June 28, 2006. Background color indicates total afternoon precipitation in mm (12-24LT), black symbols indicate grid cells excluded because of (crosses) morning precipitation (> 1mm), (triangle) topography gradients and (circles) water bodies. Events included in (excluded from) the computations are denoted by black (grey) squares. The center of each event (Lmax) is denoted with a letter; grey dots indicate Lmin. When two or more event boxes overlap, only the event with largest precipitation at Lmax is retained. Thus, a number of maxima are not interpreted as events. A total of six events are detected, four of which are included in the computation. Events “A” and “B” are not included as they include topography features or water body, respectively.
Figure C.2: For each 5° box, number of events included in the analysis (i.e., after removal of events with morning precipitation, topography or water bodies) for the CMORPH - GLEAM combination. Horizontal black lines delineate the months included in the analysis: May-September North of 23°N (upper horizontal line), November-March South of 23°S (lower horizontal line) and all months in the tropics. White boxes indicate no event, mostly related to strong topography features.

C.2 Data: detailed descriptions

C.2.1 Soil moisture datasets

We use two main soil moisture datasets based on satellite observations: surface soil moisture from AMSR-E (NASA-LPRM algorithm, Owe et al., 2008), and total evaporative stress from GLEAM (“Global Land Evaporation: the Amsterdam Methodology”, Miralles et al., 2011b).

We choose GLEAM as our primary dataset because it includes estimates of soil moisture in the whole root zone. In addition, it presents the advantage of directly providing an evaporative stress which accounts for the fact that part of the soil moisture variability (below wilting point and above the critical soil moisture level) does not affect the evaporation. It also accounts for the effect of the development of vegetation on the evaporative stress, via the inclusion of vegetation optical depth for the vegetation fractions (Miralles...
et al., 2011b). AMSR-E, on the other hand, only provides information about water in the top few cm. Note that GLEAM assimilates AMSR-E data. Here we include results from our AMSR-E-based analyses in this Supplement for comparison. Note that we did not filter AMSR-E values over densely vegetated areas such as the Amazon – these should be interpreted with caution due to the presumably poor quality of the satellite retrievals.

For each land fraction $i$, the evaporative stress in GLEAM $S_i$ is defined as a linear function between the wilting point (soil moisture level below which no water is available to plant, i.e. $S_i = 0$) and a critical soil moisture level (for and above which $S_i = S_{i,\text{max}}$, where $S_{i,\text{max}}$ is a function of vegetation optical depth). For the bare soil fraction, $S_{s,\text{max}} = 1$ (the stress for the bare soil fraction $S_s$ is also used directly in Sec. C.3). Where not stated otherwise, we use the total stress $S$ defined as the area-weighted average of individual $S_i$ values over bare soils, short and tall vegetation. Most of the day-to-day variability in $S$ is given by the soil moisture content, while vegetation controls are much slower. Therefore, we here refer to $S$ as the soil moisture stress, although it also includes information about the state of vegetation.

The original GLEAM formulation provides estimates at daily time steps. We adapt this formulation to match our specific timing requirements and thus obtain estimates at 9LT. To do so, we drive GLEAM with input data shown in Table C.1 and by aggregating all variables to a daily cycle starting and ending at 9LT, similarly to the procedure described in Guillod et al. (2014). $S$ does not include the effect of vegetation interception, but the presence of intercepted water in the morning is unlikely, as days with morning rain are removed from the analysis (Sec. C.1). To test the sensitivity of our results to

---

**Figure C.3:** Sensitivity of the metrics to the time definition of afternoon rainfall, for CMORPH - GLEAM. (top) 12-24LT, (bottom) 12-21LT. (left) Spatial metric from Fig. 4.1, (right) temporal metric from Fig. 4.2.
### Table C.1: Forcing datasets for GLEAM. Daily aggregates are computed locally to match the before-noon estimate (i.e., from 9LT on previous day to 9LT at present day to get 9LT $S$ and $S_s$).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Dataset</th>
<th>Temporal resolution</th>
<th>Spatial resolution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface soil moisture</td>
<td>NASA-LPRM (AMSR-E)</td>
<td>night-time</td>
<td>0.25°</td>
<td>Owe et al. (2008)</td>
</tr>
<tr>
<td>Vegetation optical depth</td>
<td>NASA-LPRM (AMSR-E)</td>
<td>daily</td>
<td>0.25°</td>
<td>Owe et al. (2008)</td>
</tr>
<tr>
<td>Precipitation</td>
<td>CMORPH v1.0</td>
<td>3h</td>
<td>0.25°</td>
<td>Sec. C.2.2</td>
</tr>
<tr>
<td>Precipitation</td>
<td>PERSIANN</td>
<td>3h</td>
<td>0.25°</td>
<td>Sec. C.2.2</td>
</tr>
<tr>
<td>Precipitation</td>
<td>TRMM 3B42 v7</td>
<td>3h</td>
<td>0.25°</td>
<td>Sec. C.2.2</td>
</tr>
<tr>
<td>$R_{\text{net}}$</td>
<td>CERES_SYN1deg Ed3A</td>
<td>3h</td>
<td>1°</td>
<td>Wielicki et al. (1996)</td>
</tr>
<tr>
<td>Air temperature</td>
<td>ERA-Interim</td>
<td>3h (incl. forecast)</td>
<td>0.75°</td>
<td>Dee et al. (2011)</td>
</tr>
</tbody>
</table>

The precipitation data used as input, we compute three estimates from our three precipitation datasets, which we refer to as GLEAM$_C$, GLEAM$_T$, and GLEAM$_P$ for the estimates driven by CMORPH, TRMM and PERSIANN, respectively. Results in main text refer to GLEAM$_C$.

A major difference in our input datasets compared to Guillod et al. (2014) is the use of surface radiation data from CERES (the Clouds and Earth’s Radiation Energy System, Wielicki et al., 1996) instead of the GEWEX Surface Radiation Balance (SRB) data (Stackhouse et al., 2004). These two datasets are based on multiple satellite data and they provide top-of-the-atmosphere and surface radiation fluxes globally, at a high temporal resolution (3h, e.g. Young et al., 1998). CERES is based on more recent sensors, with data starting and extending later in time, which provides longer overlap with the other products used in the analysis (CMORPH and PERSIANN in particular). In addition, CERES surface products have been shown to perform well (Kato et al., 2011, 2012, 2013).

Data gaps are filled using GPCP (for precipitation) and ERA-interim (for radiation, and for precipitation in cases where GPCP data is missing) as shown in Table C.2. All datasets are interpolated bilinearly to a 0.25° resolution prior to pre-processing. GPCP is a daily product. For each day, gaps in a 3h time step in CMORPH or PERSIANN are filled in order to match daily precipitation from GPCP. Other datasets are available at 3h resolution.

In order to mitigate the impacts of gap-filling, GLEAM 9LT values of day $i$ are masked and removed from the analysis if precipitation or $R_{\text{net}}$ is missing on day $i - 1$, or if in the 10 previous days, a gap of $n$ subsequent days in precipitation data ends between day $i - n/2$ and day $i - 1$. 

---

**Table C.1:** Forcing datasets for GLEAM. Daily aggregates are computed locally to match the before-noon estimate (i.e., from 9LT on previous day to 9LT at present day to get 9LT $S$ and $S_s$).
Variables | Dataset | Temporal resolution | Spatial resolution | Reference
---|---|---|---|---
Precipitation | GPCP\_1DD v.1.2 | daily | 1° | Huffman et al. (2001)
Precipitation (when GPCP missing) | ERA-Interim | 3h (incl. forecast) | 0.75° | Dee et al. (2011)
Net radiation | ERA-Interim | 3h (incl. forecast) | 0.75° | Dee et al. (2011)

Table C.2: Datasets used to fill GLEAM input from Table C.1. Note that temperature data from ERA-Interim does not contain any gap.

C.2.2 Precipitation datasets

We use three precipitation datasets which merge measurements from a number of satellites to produce quasi-global, consistent datasets at a high spatial (0.25°x0.25°) and temporal (3h) resolution: CMORPH (the Climate Prediction Center MORPHing method, see Joyce et al., 2004) and PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks, see Hsu et al., 1997), available from 60°S-60°N, and TRMM3B42 (from the Tropical Rainfall Measuring Mission, Huffman et al., 2007), available from 50°S-50°N and hereafter referred to as TRMM. These products provide us with at least partly independent data bases for our investigation of land-precipitation coupling.

All three products are primarily based on data from passive microwave sensor overpasses, which provide high-quality precipitation estimates but are available typically only several times a day. These are combined with data from infra-red sensors onboard geostationary satellites, which are available at a high temporal resolution over most of the globe. The products rely on different algorithms to convert the raw measurements to consistent precipitation data. TRMM estimates are also scaled to monthly rain gauge observations where available.

We choose CMORPH as our main dataset because of the more physically-based algorithm employed: It propagates passive estimates using motion-vectors derived from the sensors from geostationary satellites (see Joyce et al., 2004, for further details). We use version 1.0 of CMORPH, where the whole archive was reprocessed using a fixed algorithm and using inputs of the same versions.

TRMM uses data from the passive microwave sensors, incorporates radars on the TRMM satellites, and fills the gaps with infra-red data calibrated at the monthly time scale before scaling estimates to monthly rain gauge observations (Huffman et al., 2007). We use version 7 of the 3B42 product. PERSIANN is based on roughly the same input satellite data but uses a
neural network approach to estimate precipitation (see Hsu et al., 1997, for details).

These products have been validated and used in numerous studies (e.g. Sorooshian et al., 2000; Habib et al., 2012; Dai et al., 2007). Generally, a good agreement is found with other products (Tian and Peters-Lidard, 2010), but the quality decreases at high latitudes and over water bodies, complex terrain or coastal areas (Tian and Peters-Lidard, 2007). Fortunately, our analysis does not consider areas with complex topography and water bodies. In addition, precipitation occurrence, which is at the basis of our precipitation detection analysis, was shown to be of good quality (Tian et al., 2009).

C.3 Results with alternative datasets

A total of 12 dataset combinations are used for the analysis, combining three precipitation datasets (CMORPH, TRMM, PERSIANN) and four soil moisture datasets (GLEAM driven by our three precipitation datasets, and AMSR-E). For GLEAM soil moisture estimates, surface soil moisture stress over the bare soil fraction ($S_s$) is also included (in addition to total soil moisture stress $S$), since these may more directly relate to satellite soil moisture from AMSR-E. Results for all 12 dataset combinations are presented for spatial and temporal metrics below.

C.3.1 Spatial metric

Figure C.4 displays the spatial results from Taylor et al. (2012) as well as our results with $S$ from GLEAM (see also Fig. 4.1) and with $\Theta_{top}$ from AMSR-E, while Fig. C.5 displays the same results for $S_s$ from GLEAM. The patterns are sensitive to both the soil moisture and precipitation datasets. The signal is strongest with TRMM and weakest with PERSIANN, consistently with Taylor et al. (2012). GLEAM$_T$ also leads to a stronger signal than the other two GLEAM estimates. AMSR-E leads to a weak signal, likely because of data quality issues, while the assimilation procedure in GLEAM filters out unrealistic AMSR-E data (see also Taylor et al., 2012, who apply strict data quality filter). For GLEAM, results with $S_s$ (Fig. C.5) are smaller than with $S$ (Fig. C.4), emphasizing that $S_s$ is not always representative of the actual soil moisture stress. This might also explain part of the differences with Taylor et al. (2012) and highlights the advantage of considering the whole root zone.
Figure C.4: Spatial results (as for Fig. 4.2) for multiple soil moisture (rows) and precipitation (columns) datasets: Quantile of \( \delta_c(\Delta X') \) where \( \Delta X' = X'_{L_{\text{max}}} - X'_{L_{\text{min}}} \) and \( X' \) are anomalies of (a) surface soil moisture (\( \Theta_{\text{top}} \)) merged from AMSR-E and ASCAT (results from Taylor et al., 2012), (b-j) total soil moisture stress (\( S' \)) from GLEAM with different input precipitation data, (k-m) surface soil moisture (\( \Theta_{\text{top}} \)) from AMSR-E. Boxes with at least 25 events are displayed.
C.3. RESULTS WITH ALTERNATIVE DATASETS

Figure C.5: Same as Fig. C.4 but for surface soil moisture stress ($S_s$) from GLEAM, for various dataset combinations.
C.3.2 Temporal metric

The multi-dataset temporal results at Lmax are displayed in Fig. C.6 for $S$ from GLEAM (see also Fig. 4.2) and $\Theta_{top}$ from AMSR-E, and in Fig. C.7 for $S_s$ from GLEAM. Similarly to the spatial metrics results, the patterns exhibit some variability as a response to the choice of dataset, but the dominance of positive temporal relationships remains in all combinations. The various combinations also display some agreement in the few regions with negative relationships.

Figure C.6: Temporal results (as for Fig. 4.2) for multiple soil moisture (rows) and precipitation (columns) datasets: Quantile of $\delta_c(X'_{L_{\text{max}}})$ where $X'_{L_{\text{max}}}$ are anomalies, at Lmax, of (a-i) total soil moisture stress ($S'$) from GLEAM with different input precipitation data, (j-l) surface soil moisture ($\Theta'_{top}$) from AMSR-E. Boxes with at least 25 events are displayed.
C.3. RESULTS WITH ALTERNATIVE DATASETS

(a) CMORPH - GLEAM_C  (b) TRMM - GLEAM_C  (c) PERSIANN - GLEAM_C

(d) CMORPH - GLEAM_T  (e) TRMM - GLEAM_T  (f) PERSIANN - GLEAM_T

(g) CMORPH - GLEAM_P  (h) TRMM - GLEAM_P  (i) PERSIANN - GLEAM_P

**Figure C.7:** Same as Fig. C.6 but for surface soil moisture stress ($S'_s$) from GLEAM.
C.4 Temporal analysis from different locations

For temporal results, we have used soil moisture at Lmax. However, the overall soil moisture conditions in a larger area might be more representative of processes such as moisture recycling. Therefore, we here repeat the temporal analysis but using soil moisture averaged over Levt, the 5x5 grid cells (i.e., 1.25°) surrounding Lmax (see Sec. C.1), instead of soil moisture at Lmax. Figure C.8 displays the corresponding results for the 12 dataset combinations and can be compared to Fig. C.6. Results are roughly similar, indicating that the overall soil moisture conditions, rather than the condition at Lmax alone, might relate to afternoon precipitation. This supports the hypothesis of moisture recycling, although the various effects are difficult to explicitly disentangle, in particular the role of precipitation persistence.

![Figure C.8](image)

**Figure C.8:** Temporal results using soil moisture from Levt (5x5 grid cells surrounding Lmax) instead of Lmax (Fig. C.6), for multiple soil moisture (rows) and precipitation (columns) datasets: Quantile of $\delta_e(X'_{Levt})$ where $X'_{Levt}$ are anomalies, at Levt, of (a-i) total soil moisture stress ($S'$) from GLEAM with different input precipitation data, (j-l) surface soil moisture ($\Theta_{top}$) from AMSR-E. Boxes with at least 25 events are displayed.
C.4. TEMPORAL ANALYSIS FROM DIFFERENT LOCATIONS

Figure C.9: Temporal results using soil moisture from Lmin (location of rainfall minimum within Levt) instead of Lmax, for multiple soil moisture (rows) and precipitation (columns) datasets: Quantile of $\delta_e(X'_{L\text{min}})$ where $X'_{L\text{min}}$ are anomalies, at Lmin, of (a-i) total soil moisture stress ($S'$) from GLEAM with different input precipitation data, (j-l) surface soil moisture ($\Theta'_{\text{top}}$) from AMSR-E. Boxes with at least 25 events are displayed.

Note that moisture recycling is expected to act on longer time scales, which we do not test explicitly.

Similarly, Fig. C.9 displays results using $S'$ at Lmin, the location of rainfall minimum within Levt. Soil moisture at this location is clearly wetter for event cases than for non-event cases. This can be expected, since it is the case for $S'_{L\text{max}}$ and since the spatial metric shows that $S'_{L\text{max}}$ tends to be smaller than $S'_{L\text{min}}$ on event days, but it again highlights that the conditions before precipitation events are often wet.
C.5 Spatial coupling: the role of soil moisture heterogeneity

Spatial coupling by definition relates to soil moisture heterogeneity. Here, we investigate whether precipitation events present a preference for more heterogeneous soil moisture conditions than expected from resampling. To do so, we define heterogeneity at an event location as the spatial standard deviation of the 25 $S'$ values within $\text{Levt} (\sigma_{sp}X')$, and we compare this to the values for non-event cases, as we did with $\Delta S'$ or $S_{\text{Lmax}}$ in Fig. 4.1 and 4.2 in the main text. The corresponding results, displayed in Fig. C.10, show a clear dominance of more heterogeneous conditions than expected for precipitation events. This suggests that spatial heterogeneity might induce some precipitation events. In dry regions, $S$ is typically close to 0 and heterogeneity is

![Figure C.10: Temporal results using soil moisture heterogeneity, for multiple soil moisture (rows) and precipitation (columns) datasets. Quantile of $\delta_\sigma(\sigma_{sp}X')$ where $\sigma_{sp}X'$ is the spatial standard deviation, using the 25 grid cells within Levt, of $X'$. $X'$ are anomalies of (a-i) total soil moisture stress ($S'$) from GLEAM with different input precipitation data, (j-l) surface soil moisture ($\Theta'_{\text{top}}$) from AMSR-E. Boxes with at least 25 events are displayed.](image-url)
maximized in case of regions with wetter conditions following previous precipitation events (inducing larger value of $S$ and thereby some gradients). This suggests that precipitation events might generate following events via the spatial feedback mechanism, and thereby leading, on a large scale, to a positive feedback (Taylor et al., 2011). Contrastingly, in wet regions heterogeneity might be maximized in dry periods, leading to an overall negative feedback on the large scale, if present. This analysis highlights the possible interplays between spatial and temporal feedbacks.

C.6 Seasonality in spatial and temporal metrics

The seasonality in the metrics is presented in Fig. C.11 for CMORPH-GLEAM$_C$, and highlights some season-dependent patterns. In particular, more negative temporal relationships are found in some regions for some seasons (e.g., over the Sahel during the rainy season, JJA) compared to the seasons merged in the remaining of our analysis.

C.7 Properties of soil moisture data

Figure C.12 displays the mean and standard deviation of the evaporative stress ($S$) from GLEAM. Regions with large variability correspond to transitional regions between wet and dry climates (high and low mean $S$, respectively), where soil moisture is limiting but there is enough moisture supply to allow substantial variability.
Figure C.11: Seasonality in the coupling metrics for CMORPH precipitation and total soil moisture stress ($S'$) from GLEAM$_C$. (left) spatial metric (as in Fig. 4.1), (center) temporal metric at Lmax (as in Fig. ref:temporal), (right) temporal metric at Levt (as in Fig. C.8). Compared to the other figures, a reduced threshold of 15 events is adopted.

Figure C.12: Mean $S$ and standard deviation of $S'$ from GLEAM$_C$. GLEAM$_T$ and GLEAM$_P$ display similar patterns.
C.8 Alternative temporal metric: the simplified Triggering Feedback Strength

Our temporal metric based on precipitation event detection differs from traditional, time-series based analyses. Here, we display such a metric, the Triggering Feedback Strength from Findell et al. (2011), which quantifies the relationship between before-noon EF and afternoon precipitation occurrence as

$$ TFS = \sigma_{EF} \frac{\partial \Gamma(r)}{\partial EF}, $$

where \( \Gamma(r) \) is the probability of afternoon rainfall (\( r > 1 \)mm) and EF is before-noon (9-12LT) evaporative fraction (\( EF = \lambda E/(H + \lambda E) \)). We replace EF by \( S \) and we use anomalies from the seasonal cycle (\( S' \)). In addition, we scale TFS by the mean afternoon precipitation occurrence (\( \Gamma(r) \)) to allow for comparison between regions with different precipitation regimes. Thus, we define a simplified Triggering Feedback Strength,

$$ sTFS(S') = \frac{\sigma_{S'}}{\Gamma(r)} \frac{\partial \Gamma(r)}{\partial S'}, $$

where \( \sigma_{S'} \) is the standard deviation of \( S' \) (using 9-12LT values from each day), and \( \Gamma(r) \) is the probability of afternoon rainfall (\( r > 1 \)mm for the 12-24LT time period). The numeric computation uses two bins, and significance is tested by means of 1000 bootstraps samples as in Guillod et al. (2014). In addition, the original computation from Findell et al. (2011), binned on variables relating to atmospheric humidity and stability, is replaced by the simpler computation from Eq. C.1. Days with morning precipitation exceeding 1mm in any neighbouring grid cells in a box of 1.25° surrounding each grid cell are excluded from the computation.

Fig. C.13 displays \( sTFS(S') \) and its p-values relative to the null distribution obtained from bootstrapping. Note that no topography- or water-related filter is applied. Comparing these results with our temporal metric

![Fig. C.13: Scaled, simplified Triggering Feedback Strength (sTFS) for CMORPH - GLEAMC. (a) sTFS(S'), (b) p-value (significance). Horizontal black lines delineate the months included in the analysis: May-September North of 23°N (upper horizontal line), November-March South of 23°S (lower horizontal line) and all months in the tropics.](image-url)
(a) sAFS($S'$)  
(b) p-value (significance)

**Figure C.14:** Scaled, simplified Amplification Feedback Strength (sAFS) for CMORPH - GLEAM$_C$. (a) sAFS($S'$), (b) p-value (significance). Horizontal black lines delineate the months included in the analysis: May-September North of 23°N (upper horizontal line), November-March South of 23°S (lower horizontal line) and all months in the tropics.

(e.g., Fig. 4.2 from the main text), it is interesting to note that the significance (p-values) depicts similar regions of positive and negative temporal relationships.

While our methodology focuses on the impact of soil moisture on precipitation occurrence, Findell et al. (2011) also introduced a metric quantifying the relationship between before-noon EF and afternoon precipitation amounts: The Amplification Feedback Strength (AFS), restricted to days with afternoon rain (> 1mm). Figure C.14 displays sAFS, a simplified AFS formulation similar to what sTFS is to TFS. No strong relationship is found anywhere, suggesting that the local impact of soil moisture on rainfall amounts is negligible and consistently with the findings from Findell et al. (2011) over North America. Nonetheless, we note that the accuracy of rainfall amounts has been shown to be low relative to the accuracy of rainfall occurrence (Tian et al., 2009), which may prevent the exclusion of such local impacts on precipitation amounts.
Bibliography


assimilation system. Quarterly Journal of the Royal Meteorological Society 137 (656), 553–597.


URL www.cosmomodel.org


Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reich, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., Zeng, N., Jul. 2006. Climate-carbon cycle feedback analysis: Results from the c4mip model intercomparison. Journal of Climate 19 (14), 3337–3353.


URL http://ipcc-wg2.gov/SREX/


Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., van Beek, L. P. H., Uhlenbrook, S., 2014. Comparison of different evaporation...


I would like to thank a number of people for their support during the past three years.

First of all, I would like to thank my PhD advisor and supervisor, Sonia. Not only did she give me the opportunity to do my PhD in her group, but she has also been very supportive and motivating all these years. She is a very talented and recognized scientist and it has been a great pleasure to work under her supervision. I am also indebted to her for the freedom she gave me, the opportunities to go to great scientific conferences, as well as for introducing me to many scientists around the world. Merci, Sonia!

I am very thankful to Boris, my direct supervisor, with whom I enjoyed to work with a lot. His clear guidance, straightforward thinking and patience clearly contributed, not only to make me progress in the right direction, but also to avoid getting lost in the complex topic I worked on. He also supported me during the most difficult parts of my PhD and provided extensive and valuable feedback on my ideas and the paper drafts. I wish him all the best in his new pursuits. Danke Boris!

I would also like to thank Chris, for great discussions at various conferences but even more for accepting to take the seat of the external examiner for my defense. Given his ground–breaking contributions to the field of land-precipitation coupling, it is a great pleasure to have him there for the last hurdle.

Many other people contributed to my work, and I thank them all for their contributions. Special thanks to Pierre Gentine for my stay at Columbia University, as well as to Diego and Ryan for their interest and for motivating discussions. Thanks also to Urs and Dani for the wonderful IT support, which made things so much easier!

I’d also like to thank Edouard, who supervised my post-diploma work and the first part of this thesis. Working with him was great and it was probably a decisive element in my choice to pursue a PhD. But most of all, I’d like to thank him for his friendship and the dinners at Zähringer, beers in BQM, and Sunday’s walks on the Üetliberg!

Some people deserve special thanks: Quentin (Mr. T.) and Paulo made
the first 2 years of my thesis and my life in Zurich a real good time. Merci les gars! Thanks also to Nathalie for the coffee breaks and chats, and Brigitte for wonderful discussions and for hosting me a few weeks.

The great atmosphere within the landclim group made the last three years very enjoyable. Particular thanks to Ruth for being always in for a coffee and to Heidi for the baby-sketch (I’ll remember that one for a long time!). Thanks to all landclim members!

I’d also like to thank many colleagues and friends at the institute for the wonderful time I have spent here, including Joeri, Omar, Maria, Jan, Erica, Franziska, Dave, and Sven.

Some older friends also deserve to be named here. The SIE friends at EPFL (Pierre, Simon, Tristan, David and many others), great flatmates (Yann, Raf, Manu, Mauro), the IIT Madras team (Danielle, Anna, Reto, Olivier, Björn, Fanny, Ben and others), as well as Guillaumarc. I am also very thankful to my family for their support all these years.

Finally, I would like to thank Amanda for her support and love.

Benoît Guillod, December 2013
Curriculum Vitae

Benoît Guillod
Institute for Atmospheric and Climate Science
ETH Zürich
Universitätstrasse 16
8092 Zürich, Switzerland

Personal
Swiss citizen, born 04/04/1986 in Lausanne, Switzerland

Academic Background

since 10/2010 ETH Zurich, Switzerland
PhD student/Research Assistant
Institute for Atmospheric and Climate Sciences, Land-Climate Dynamics Group
Topic: Observation- and modeling-based investigations of land-precipitation coupling
Supervisors: Prof. Dr. Sonia I. Seneviratne and Dr. Boris Orlowsky

09/2007–01/2010 ETH Zurich, Switzerland
Master of Science ETH in environmental sciences
Major in “Atmosphere and Climate”
Master thesis: Footprint Modeling for Heavy Particles
Supervised by PD. Dr. Mathias Rotach
Minors in “Physical Glaciology” and in “Sustainable Energy Use”

08/2006–06/2007 Indian Institute of Technology (IIT) Madras, Chennai, India
Exchange year
10/2004–07/2007  **EPFL, Lausanne, Switzerland**  
**Bachelor of Science in Environmental Sciences and Engineering**

08/2001–07/2004  **Lycée Blaise-Cendrars, La Chaux-de-Fonds, Switzerland**  
**Matura** (Swiss high school degree)  
Major in Physics and Mathematics  
Minor in Law and Economics

**Work experience**

06/2010–09/2010  Research Internship at ETH Zurich, Institute for Atmospheric and Climate Sciences, Land-Climate Interactions Group (Prof. Dr. Sonia I. Seneviratne)


09/2008–02/2009  Internship at Community, Habitat and Environmental Management (CHEMA), Tanzania, under the aegis of the Swiss Agency for Development and Cooperation (SDC), for the project “CHEMA Rocket Stove”

**Teaching**

Teaching Assistant:  “Land-Climate Interactions” (MSc course, fall 2011, 2012)  
“Climatological and Hydrological Field Work”  
(MSc course, spring 2012)  
“Numerical Modeling of Weather and Climate”  
(MSc course, spring 2011)  
“Atmospheric physics lab work” (MSc course, spring 2011)

**Supervision**

MSc thesis: Mathias Hauser (10/2012–03/2013)  
“Sensitivity of Soil Moisture and Ground Thermal Regimes to Soil Organic Carbon in the Community Land Model”
Publications


Oral presentations

4th iLEAPS Science Conference, Nanjing, China, 2014. “Soil moisture-precipitation coupling: a comparison of temporal versus spatial perspectives”.

American Geophysical Union Fall Meeting, San Francisco (CA), USA, 2013. “Land surface controls on afternoon precipitation diagnosed from observational data: Uncertainties, confounding factors and the possible role of interception storage”.

European Geoscience Union General Assembly, Vienna, Austria, 2013. “Soil moisture–precipitation coupling: observational analysis of the impact on precipitation occurrence in North America”.

GLASS panel meeting (Global Land/Atmosphere System Study, GEWEX), Boulder (CO), USA, 2012. “GLACE-CMIP5” (work from Sonia I. Seneviratne).

GLASS panel meeting (Global Land/Atmosphere System Study, GEWEX), Boulder (CO), USA, 2012. “Investigation of LoCo in observations: issues and ideas”.

179
Pan-GASS Workshop (1st Pan-Global Atmosphere System Study Conference), Boulder (CO), USA, 2012. “Soil moisture–precipitation coupling using observations from FLUXNET, precipitation radar and reanalysis”.


COSMO-CLM user meeting, Langen, Germany, 2011. “Comparison of FAO versus JRC soil map in COSMO-CLM”.

**Poster presentations**


Pan-GASS Workshop (1st Pan-Global Atmosphere System Study Conference), Boulder (CO), USA, 2012. “Soil moisture–precipitation coupling using observations from FLUXNET, precipitation radar and reanalysis”.


Swiss Global Change day, Bern, Switzerland, 2012. “Impact of soil map specifications for European climate simulations”.


European Geoscience Union General Assembly, Vienna, Austria, 2011. “Observation-based investigation of land-precipitation feedback”.


180