Jessica Keune

Integrated terrestrial simulations at the continental scale:

Impact of groundwater dynamics and human water use on groundwater-to-atmosphere feedbacks during the European heatwave in 2003

Jessica Keune

INTEGRATED TERRESTRIAL SIMULATIONS AT THE CONTINENTAL SCALE:

IMPACT OF GROUNDWATER DYNAMICS AND HUMAN WATER USE ON GROUNDWATER-TO-ATMOSPHERE FEEDBACKS DURING THE EUROPEAN HEATWAVE IN 2003

Integrated terrestrial simulations at the continental scale: Impact of groundwater dynamics and human water use on groundwater-to-atmosphere feedbacks during the European heatwave in 2003

vorgelegt von
M. Sc. Jessica Keune
aus
Troisdorf

Bonn, Dezember, 2018

Anschrift des Verfassers: Address of the author:

Jessica Keune
Institut für Geowissenschaften und Meteorologie der
Universität Bonn
Auf dem Hügel 20
D-53121 Bonn

Abstract

The availability of freshwater under current and future climatic conditions remains one of the main research topics that requires an improved understanding of the terrestrial system and anthropogenic impacts. Yet, hydrologic and atmospheric studies are commonly performed in a disjunct fashion even though hydrologic and atmospheric processes are intrinsically connected through the terrestrial water, energy, momentum and matter cycles. In particular, current continental-scale process and assessment studies do not close these cycles, neglect groundwater dynamics and miss human impacts related to water use.

Continental-scale, coupled modeling systems are required to improve process representations of the terrestrial cycles and their feedbacks. In this thesis, a fully-coupled, groundwater-land surface-atmosphere modeling system, the Terrestrial Systems Modeling Platform TerrSysMP, is setup over Europe, in order to simulate the full terrestrial water cycle from groundwater across the land surface into the atmosphere. The model setup is designed to perform sensitivity studies, which assess the impact of (i) groundwater dynamics and (ii) human water use on the continental-scale terrestrial water cycle during the heatwave year in 2003. Beyond that, the simulations are used to define (iii) the associated added value of incorporating human water use.

First, over continental Europe, the impact of groundwater dynamics on associated land-atmosphere feedbacks during the peak of the heatwave is evaluated. Results illustrate the potential of groundwater dynamics to dampen heat extremes, especially in regions of shallow water tables, and uncover potential deficiencies of current assessment studies. Ensemble simulations are performed and address the intrinsic uncertainty of land-atmosphere feedbacks to subsurface characteristics, such as soil texture and the incorporation of an underlying geology. This study shows that the groundwater representation induces variability across the compartments of the terrestrial system, with a decreasing impact from the subsurface over the land surface towards the atmosphere. The impact of groundwater is strongly dependent on the prescribed subsurface characteristics, i.e. the distribution of hydrofacies and soil hydraulic parameters.

Second, the incorporation of human water use related to groundwater abstraction and irrigation, and its impact on water availability and fluxes, is evaluated. Four water use scenarios are constructed to account for the uncertainty of human water use and the associated land-atmosphere feedbacks. The simulations indicate that atmospheric feedbacks to human water use may contribute systematically to continental drying and wetting in Southern Europe. Moreover, the simulated land-atmosphere feedbacks exceed net values

of human water use (i.e. irrigation minus groundwater pumping) at the watershed scale, and significantly affect atmospheric moisture transport across watersheds. These results emphasize that effects of local human water use on water availability and sustainability may be of global importance.

Finally, the added value of incorporating human water use in terrestrial simulations and predictions of evapotranspiration and precipitation is examined. Multiple observational data sets are used to assess observation uncertainty and model skills. The results illustrate model deficiencies, which mainly arise from the simulation of precipitation and an overestimation of extreme precipitation events. Observational uncertainties, however, exceed the impact of human water use on precipitation and evapotranspiration at the watershed scale, and impede model validation. At the local scale, human water use significantly affects daily precipitation and evapotranspiration, and potentially improves the model skill through an improved simulation of local precipitation.

Integrated feedbacks from groundwater to the atmosphere, and through the incorporation of human water use, are not yet considered in regional climate simulations, reanalyses and water resource assessment studies, but constitute important processes and additional uncertainties in simulations of the terrestrial water cycle. Therefore, this thesis emphasizes the need and advantages of continental-scale, coupled modeling systems, to advance impact studies across compartments of the Earth systems.

Contents

Chapter 1

Introduction

1.1 Background

One of the major socioeconomic challenges is to secure fresh water availability to humans, under present and future environmental conditions. While about 70% of the Earth's surface is covered by water, only about 2.5% of the world's water is fresh water that is available to humans (Oki and Kanae, 2006). Currently, more than 4 billion people worldwide suffer water scarcity at least once a year during e.g. droughts and heatwaves (Mekonnen and Hoekstra, 2016; Vörösmarty et al., 2000). Water contamination, population growth, urbanization, industrialization and agricultural expansion, but also climate change along with a potential intensification of the water cycle with more frequent and intense droughts (Huntington, 2006) cause an increasing risk of water scarcity today and in the future. Especially the use of water for agricultural production, to e.g., maintain food supply, accounts for more than 90% of the global water use footprint (Hoekstra and Mekonnen, 2012). Overall, this may cause an unsustainable use of water resources (Jaramillo and Destouni, 2015) and enhance continental drying (Milly and Dunne, 2016), which ultimately increases the threat for socio-economic wellbeing and security.

Unfortunately, observations across climate subsystems (e.g., from the pedosphere to the biosphere and to the atmosphere) are scarce and do not lend themselves for an accurate monitoring of the terrestrial cycles due to observational errors and uncertainties, and limited availability across spatial and temporal scales (McCabe et al., 2017; Wood et al., 2011). To a large extent, our knowledge of the terrestrial water cycle and its uncertainties relies on numerical models, which simulate the terrestrial system and allow to perform predictions of future climate and water availability. However, numerical studies commonly address the unidirectional cause-effect chain (Rast et al., 2014), i.e., the effect of climate change on terrestrial hydrology, human water use and water resources, but neglect the two-way feedback between the different compartments and the anthropogenic activities. Vice versa, land surface models employed in current regional or global climate models often do not close the terrestrial water cycle, apply prescribed atmospheric forcing, parameterize groundwater and hence cannot be used to study the two-way feedbacks between the atmospheric circulation and water resources.

This thesis aims to provide insight into two-way feedbacks and uncertainties of the continental-scale terrestrial water cycle, and attempts to improve our understanding of regional model deficiencies related to groundwater dynamics and human water use, which may contribute to an improved simulation of water availability. This chapter presents an overview and discussion of state-of-the-art research on feedbacks between subsurface, land surface and atmospheric processes, and the impact of human water use on the terrestrial water cycle. Subsequently, the research objectives of this thesis are defined and the outline of the different chapters is presented.

1.2   Terrestrial feedbacks

1.2.1   Land-atmosphere feedbacks

Land-atmosphere feedbacks describe the interaction between the land surface and the atmosphere through the exchange of mass, energy and momentum. Moisture, commonly considered as soil moisture at the land surface, and heat have been shown to strongly influence the atmospheric boundary layer (ABL) through the partitioning of the land surface fluxes. The incoming net radiation is partitioned into the sensible heat, the latent heat and the ground heat flux, with the latter being comparably small. Two main feedback mechanisms have been identified, namely, the soil moisture-temperature and the soil moisture-precipitation feedback, and describe the land surface's influence on climate variability (Seneviratne et al., 2006; Fig. 1.1). The soil moisture-temperature feedback is commonly depicted as a local feedback loop and formulated in a simplistic way, neglecting other feedback processes, e.g., with clouds. A decrease of soil moisture associated with increasing temperatures, leads to a decrease of evapotranspiration from the land surface, which may further intensify near-surface (and lower atmospheric) temperatures (Fig. 1.1a). This feedback loop is not entirely positive, because a temperature increase is expected to lead to an evapotranspiration increase rather than a decrease, which might furthermore alleviate the initially induced soil moisture decrease.

The main physical mechanism behind this feedback is the partitioning of the land surface fluxes, in particular the latent heat flux associated with evapotranspiration, and the sensible heat flux, as a function of soil moisture content. Studies distinguish between two land-atmosphere feedback regimes: (i) soil moisture-limited and (ii) energy-limited (Koster et al., 2004; Seneviratne et al., 2006). The latter describes a scenario in which soil moisture is not limited and evapotranspiration is determined by the net radiation and hence describes the radiation control on land surface processes (similar to the precipitation control in case of the water balance). On the contrary, during a soil moisture limitation, the land surface controls the partitioning of the surface fluxes, and the amount of water being evaporated is a function of soil moisture content until a critical threshold (e.g., the wilting point) is reached (Seneviratne et al., 2010).

The soil moisture-temperature feedback is, due to its local impact, the most commonly studied influence of the land surface on the atmosphere (e.g., Dirmeyer, 2011; Seneviratne et al., 2006; Seneviratne et al., 2010; Miralles et al., 2012). The feedback is often associated with extreme heat and drought (e.g., Hirschi et al., 2010; Jaeger and Seneviratne, 2010) and has been shown to be one of the main drivers for heatwaves, such as the European heatwave in 2003 (e.g., Miralles et al. 2014; Fischer et al., 2007; Zaitchik et al., 2006; Stéfanon et al., 2013; Fennessy and Kinter, 2011).

Especially heatwaves are associated with precipitation deficits (Mueller and Seneviratne, 2012; Vautard et al., 2007). Yet, the soil moisture-precipitation feedback (Fig. 1.1b) is not necessarily local due to long-range interactions in time and space, which makes the feedback more difficult to detect using both, observations and simulations. A simplistic description of the feedback is as follows: changes of soil moisture and evapotranspiration further impact precipitation. However, the sign of this feedback is not clear as it strongly depends on the
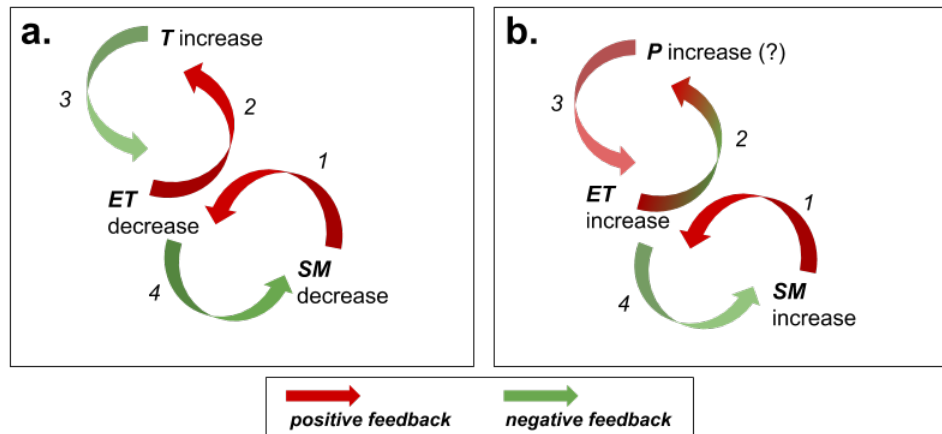
Figure 1.1: Sketch of (a) the soil moisture (SM)-temperature (T) and (b) the SM-precipitation (P) feedback after Seneviratne et al. (2010). Numbers indicate process steps of the feedback loop (an expanded explanation can be found in the text). Red arrows indicate positive feedback loops, i.e., an intensification of the feedbacks. Green arrows indicate negative feedback loops, i.e., an alleviation of the feedback. Uncertain feedbacks are indicated by red-green arrows. Note that feedback mechanisms are highly simplified and especially precipitation feedbacks are not necessarily local. Redrawn after Seneviratne et al. (2010).

atmospheric state. Hence, a soil moisture decrease may ultimately inhibit, but can also initiate precipitation. Evidence for soil moisture-precipitation feedbacks is mostly limited to coherence and covariance of dry soils with (afternoon) rainfall (e.g., Taylor et al., 2012; Eltahir, 1998; Findell et al., 2011; Guillod et al., 2015; Guillod et al., 2014), or arises through the comparison of simulation experiments (Koster et al., 2003; Koster et al., 2004).

1.2.2 Groundwater-atmosphere feedbacks

The hydraulic connection of soil moisture to groundwater has often been neglected in the description of land-atmosphere feedbacks and regional to global climate assessments. However, studies indicate that groundwater dynamics largely determine evapotranspiration via lateral subsurface flow and a dynamic groundwater table, not only at catchment and regional scales (e.g., Chen and Hu, 2004; Yeh and Eltahir, 2005; Kollet and Maxwell, 2008; Maxwell and Kollet, 2008; Ferguson and Maxwell, 2010; Lam et al., 2011; Soylu et al., 2011; Rahman et al., 2014), but also at the continental and global scale (e.g., Niu et al., 2007; Koirala et al., 2017; Maxwell and Condon, 2016; Ji et al., 2017). This connection is amplified through vegetation, which may adjust the rooting depths towards the water table (Fan et al., 2017). Groundwater is hence a source of water for plants not only in semi-arid to arid regions, and actively contributes to the formation of landscape, land cover and regional climate zones (Fan, 2015).

The connection of land surface fluxes with groundwater further influences atmospheric processes via groundwater-land surface-atmosphere feedbacks. These feedbacks have been assessed in sensitivity studies using different bottom boundary conditions (mimicking e.g., an impermeable bottom and a constant water table depth) or an improved hydrology in land surface models (e.g., Johnson et al., 1993; Chen and Dudhia, 2001; Campoy et al., 2013), or

simplified atmospheric conditions and models with a more complex hydrologic model (e.g., Quinn et al., 1995; York et al., 2002). There exist a number of studies, which employed integrated soil-vegetation-atmosphere models to study the effects of groundwater dynamics on land-atmosphere feedbacks (Walko et al., 2000; Seuffert et al., 2002; Miguez-Macho and Fan, 2012; Anyah et al., 2008; Maxwell et al., 2007, 2011; Shrestha et al., 2014; Butts et al., 2014; Larsen et al., 2014; Barlage et al., 2015; Davison et al., 2018). Simulating the two-way feedbacks between groundwater and atmosphere, these studies indicate that a dynamic representation of groundwater is especially important under moisture-limited conditions and can significantly affect near-surface conditions, boundary layer processes and precipitation (Rahman et al., 2015; Gilbert et al., 2017), and amplify temperature extremes (Davison et al., 2015; Fig. 1.1). It is noted, however, that most of these studies are set up at the catchment scale. Just recently, Barlage et al. (2015) showed that the representation of groundwater representations may also have significant effects at the continental scale, even with a relatively coarse grid-spacing of 30 km over the continental United States and during a simulation period of 6 months.

1.2.3 Human water use feedbacks

Changes in precipitation over the continental United States have been shown to be related to the development of groundwater pumping and intensified irrigation practices during the last century (Alter et al., 2015; DeAngelis et al., 2010; Moore and Rojstaczer, 2001), thereby exemplary illustrating human water use feedbacks with the atmosphere. However, feedback pathways and magnitude remain uncertain and constitute a focus of current research activities.

Due to the varying availability of water in space and time, humans pump groundwater to supplement the demand. Extensive groundwater pumping leads to declining water tables in aquifers around the world (e.g., Famiglietti, 2014; Long et al., 2013; Rodell et al., 2009; Scanlon et al., 2012), thereby putting water resources at risk. Studies showed that regions with low recharge rates are particularly susceptible to groundwater depletion (Leng et al., 2014), and that groundwater pumping for irrigation amplifies the annual stream flow cycle (Condon and Maxwell, 2014a). Moreover, terrestrial variability induced by human water use is approximately on the same order of magnitude as climate change variability (Ferguson and Maxwell, 2012), and large-scale, simplified hydrologic models tend to underestimate water storage anomalies as evaluated with satellite observations (Scanlon et al., 2018). In the future, an increasing demand of water may intensify hydrological droughts (Wada et al., 2013). However, in analogy to the above prescribed limitations, these studies were either carried out with simplified, conceptual hydrologic models, or used prescribed atmospheric forcing, thereby neglecting the full, physics-based feedback pathway from groundwater to the atmosphere.

On the contrary, atmospheric impact studies focus on the influence of irrigation on land-atmosphere feedbacks and precipitation. All studies agree on a consistent cooling of the near-surface atmosphere induced by irrigation, but with varying magnitude (e.g., Haddeland et al., 2006a; Kueppers et al., 2007; Lobell et al., 2008; Sacks et al., 2009; Puma

and Cook, 2010; Kueppers and Snyder, 2012; Lu and Kueppers, 2015; Cook et al., 2015; Thiery et al., 2017). Local to regional precipitation response to irrigation remains uncertain (e.g., Harding and Snyder, 2012; Kueppers and Snyder, 2012; Moore and Rojstaczer, 2001; Pei et al., 2016). The precipitation response to irrigation depends on the atmospheric state and requires certain atmospheric conditions, such as moisture convergence, which allows the additional moisture to ascend towards the clouds (Barnston and Schickedanz, 1984). In contrast to the near-surface atmospheric cooling, effects of irrigation on precipitation are not necessarily local and can bypass continents (de Vrese et al., 2016), further impacting remotely river discharge (Wang-Erlandsson et al., 2017). Yet, simulations of irrigation practices in atmospheric models are often implemented in a simplified fashion and do not account for the origin of water for irrigation, i.e., irrigation is usually mimicked through soil moisture at field capacity or a constant potential evaporation flux, both arising from unknown sources. Improved implementations of human water use incorporate the water stress of the plants and directly remove the applied water amounts from runoff estimates (Thiery et al., 2017), while neglecting other water sources and their feedbacks. Moreover, these studies cannot draw any conclusion on the impact of irrigation practices on water resources. Recent approaches incorporate the effect of groundwater pumping in land surface models (Pokhrel et al., 2015), but do not account for the physical feedback pathways.

1.2.4   Towards Earth System Models

The previously described interactions and model deficiencies have been identified, and approaches towards improved representations of terrestrial hydrology in atmospheric models (Hazenberg et al., 2016) are underway. Often, however, conceptual hydrologic models are used, but require a calibration using full-physics models (Clark et al., 2015). Moreover, conceptual hydrologic models can only parameterize the human impact through, e.g., groundwater pumping on terrestrial hydrology (Nazemi and Wheater, 2015; Wada et al., 2017).

An integrated view of the terrestrial system is required to assess two-way feedbacks across the terrestrial hydrologic cycle and to quantify uncertainties in current model simulations to advance our knowledge of water availability and sustainability. That includes an impact assessment of groundwater dynamics and human water use on atmospheric processes (weather and climate), and, vice versa, the feedback of climate change on terrestrial hydrology including water resources. Therefore, individual component models of the terrestrial system have been coupled to resemble the Earth system and simulate processes across the climate subsystems, maintaining the conservation of mass, energy and momentum. The resulting multi-component models are typically called Earth System Models (ESM, Giorgi, 1995). Specifically, the coupling of more detailed and physics-based hydrologic models to land surface- and atmospheric models affords studies across the soil-vegetation-atmosphere system (Simmer et al., 2015).

However, the spatial extent of the above mentioned simulations has been mostly limited to the catchment scale due to the computational demand of atmospheric and physics-based
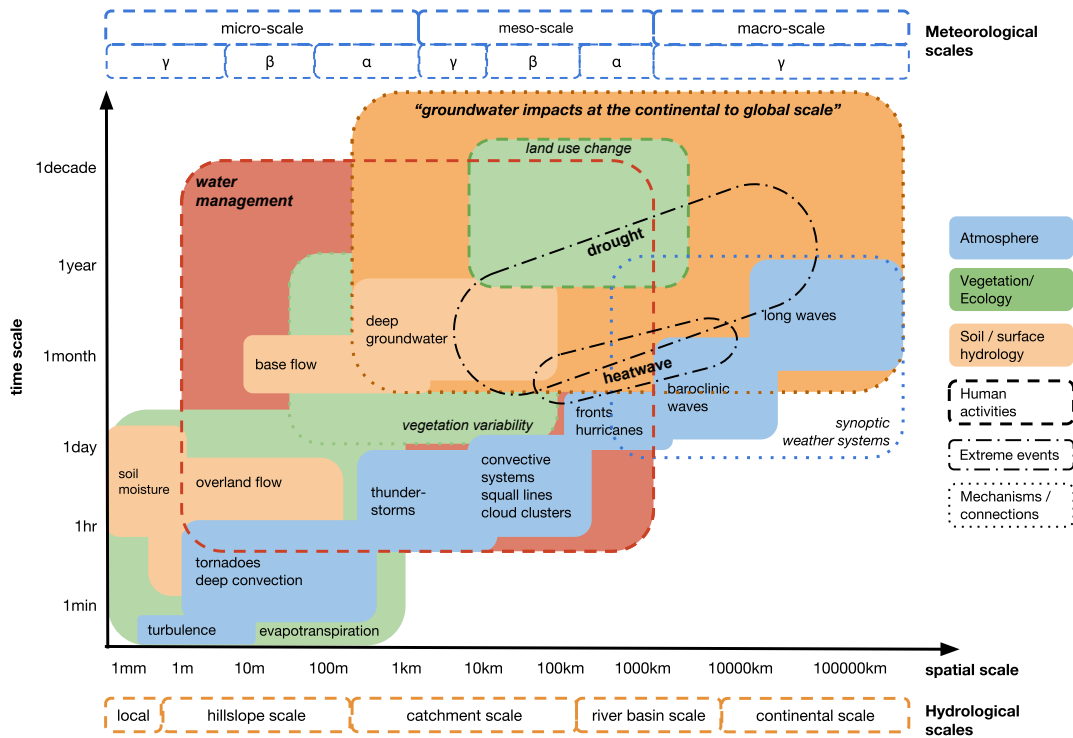
Figure 1.2: Sketch of the spatial and temporal scales for atmospheric (blue), vegetation/ecology-related (green) and hydrologic land surface-subsurface processes (orange), including typical time scales of human interaction (land use change; water management) and hydro-meteorological extremes, such as droughts and heatwaves. Atmospheric and hydrologic scales are based on Orlanski (1975) and Brutsaert (2005), respectively.

groundwater simulations. Analogously, in order to limit the computational burden, groundwater has been parameterized applying simplifying assumptions, which have not been evaluated comprehensively (Clark et al., 2015). This thesis demonstrates, however, that recent advances in scientific computer software and hardware technologies allow to implement parallel, fully-coupled soil-vegetation-atmosphere models to perform simulations in reasonable time, and to increase complexity, resolution, spatial and temporal extent, simultaneously.

There are several scientific rationales for the application of soil-vegetation-atmosphere models beyond the catchment scale. Climate subsystems act and interact at various spatial and temporal scales (Fig. 1.2). Synoptic conditions often determine the atmospheric water flows and the strength of the influence of the land surface on atmospheric processes. Vice versa, in situations where the land surface significantly affects atmospheric processes, land-atmosphere feedbacks are not limited to the catchment scale and can have remote effects through, e.g., changes of the atmospheric water vapor transport from the atmospheric micro- to the macro-scale, thereby motivating regional to global scale atmospheric simulations. The same rationale applies to large-scale hydrologic models. Here, the underlying hypothesis is that there exists a groundwater impact at the continental scale through, e.g., lateral subsurface flow and emerging regional to continental scale gradients and spatial patterns of water table depth, which influence the land surface and the atmosphere (Fan, 2015). Studies emphasize the importance of lateral subsurface flow for establishing realistic water budgets in climate models (Krakauer et al., 2014; Fan, 2015), which is especially important for the

assessment of water resources. Similarly, irrigation has been shown to impact precipitation beyond the catchment scale and from one continent to another (de Vrese et al., 2016). In summary, these findings illustrate the need of continental-scale simulations of the terrestrial system, including two-way groundwater-to-atmosphere feedbacks.

1.3   Objectives and outline

This thesis aims to improve the fundamental understanding of the processes and uncertainties of the terrestrial water cycle at the continental scale related to the groundwater representation and human water use. Therefore, three main research objectives are formulated. The first objective of this thesis is to identify and assess the impact of groundwater dynamics on land-atmosphere feedbacks at the continental scale, using a coupled soil-vegetation-atmosphere modeling system. The second objective focuses on an assessment of the contribution of human water use induced atmospheric feedbacks to water availability in this modeling system. The third objective is to evaluate the latter simulations with observations and identify a potential improvement through the incorporation of human water use.

To address these objectives, this thesis formulates three major research questions:

1. How does groundwater contribute to land-atmosphere feedbacks at the continental scale and to which extent might groundwater buffer extreme heat, e.g. during droughts and heatwaves?

2. Do atmospheric feedbacks to human water use affect water resources, and if so, to which extent might they contribute to continental drying (or wetting)?

3. Does the incorporation of human water use in a continental-scale modeling system improve the simulation of evapotranspiration and precipitation?

This thesis is structured as follows. Chapter 2 introduces the Terrestrial Systems Modeling Platform (TerrSysMP; Shrestha et al. 2014; Gasper et al., 2014) and its components, which are used in this thesis to answer the above described research questions. It gives a brief overview of the setup of this modeling system over Europe, the workflow for fully-coupled simulations, and describes the challenges associated with an integrated, soil-vegetation-atmosphere modeling system.

Subsequently, Chapter 3 addresses the first research question and evaluates the impact of groundwater dynamics on land-atmosphere feedbacks at the continental scale during the European heatwave in 2003. The fully-coupled soil-vegetation-atmosphere modeling system is setup over the Europe with two different groundwater physics: (i) a typical 1D free drainage approach, which is often applied in land surface models; and (ii) a 3D full physics groundwater model. The differences between the two approaches are evaluated considering additional uncertainties of the subsurface parameters and the initial condition. This chapter focuses on the uncertainty of soil moisture-temperature feedbacks in regional climate simulations with respect to the groundwater representation. The effects of different groundwater representations on the magnitude of land-atmosphere feedbacks are assessed,

and differences of key atmospheric variables, such as temperature and convective available potential energy are illustrated. A statistical model evaluates the impact of all factors (groundwater representation, subsurface characteristics, initial condition) on the variability of water cycle variables across the terrestrial system. This substantiates a discussion of groundwater-induced variability considering the background variability, which is mainly determined by the atmospheric circulation.

Chapter 4 focuses on the second research question and evaluates the impact of human water use on the continental sink of water, i.e., the difference between evapotranspiration and precipitation, within a fully-coupled soil-vegetation-atmosphere modeling system over Europe. Here, two data sets of realistic estimates of groundwater pumping and irrigation, along with two human water use schedules, are used to assess the uncertainty of the simulated land-atmosphere feedbacks to human water use. The consistency of the feedback signal is analyzed at various temporal and spatial scales, i.e., from the grid point to the catchment and from the seasonal to the annual time scale. The integrated modeling approach allows to identify a relationship between atmospheric feedbacks to human water use and water resources. The magnitude of the human water use-induced atmospheric feedbacks is compared to the direct effects of pumping and irrigation, allowing for an objective interpretation of the significance.

Chapter 5 comprises an extensive evaluation of the simulations from Chapter 4 with observations at the land-atmosphere interface, addressing the third research question in this thesis. The focus is set on the added value of incorporating human water use on the simulation of the terrestrial water cycle. First, the simulation accuracy at the watershed scale is evaluated. The use of multiple observational data sets across temporal scales allows to assess observational uncertainty. In addition to a bias assessment, probabilistic verification measures validate the model capabilities to simulate timing and location of precipitation events. Subsequently, the added value of incorporating human water use on the simulation of precipitation and evapotranspiration is assessed. Typical verification measures, such as the bias and the mean squared error are analyzed, and changes due to the incorporation of human water use, are discussed. Improvements at the watershed scale are interpreted considering observational uncertainties. Local improvements are evaluated using daily, (co-located) *in-situ* observations of precipitation and evapotranspiration, and differences between regions are discussed.

Finally, Chapter 6 summarizes the answers obtained to the above formulated research questions. It embraces a brief résumé of the results of each study, discusses assumptions and potential limitations, and provides recommendations for future research.

Chapter 2

The Terrestrial Systems Modeling Platform - TerrSysMP

2.1 TerrSysMP

The Terrestrial Systems Modeling Platform, TerrSysMP, originated out of a model development effort of the Transregional Collaborative Research Centre TR32[1] *Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modeling and Data Assimilation* (Simmer et al., 2015) aiming at an integrated simulation approach to close the coupled terrestrial water and energy cycle from groundwater across the land surface into the atmosphere. The initial version was published by Shrestha et al. (2014) consisting of three individual model components that are the numerical weather prediction system COSMO of the German Meteorological Service (*Deutscher Wetterdienst*, DWD); the land surface model Community Land Model Version 3.5 of the National Center for Atmospheric Research (NCAR), USA, and the surface-subsurface flow model ParFlow, a joint development effort of the Lawrence Livermore National Laboratory, Colorado School of Mines, USA; and the Research Centre Jülich, Germany. Initially, the individual model component of TerrSysMP were coupled using the Ocean Sea Ice Soil (OASIS3) coupler. In an ensuing step, the coupling was extended to OASIS3-MCT, using the Model Coupling Toolkit (Warner et al., 2008), which affords better parallel performance in case of very large models and massively parallel supercomputer environments based on the Multiple-Program Multiple-Data (MPMD) programming approach (Gasper et al., 2014).

The following subsections briefly introduce the individual model components and describe the physical coupling based on the exchange of states and fluxes at the subsurface-land surface and land surface-atmosphere interfaces. For a detailed description of the model components, the reader is referred to the listed publications.

2.1.1 COSMO

The atmospheric model COSMO (COnsortium for Small-Scale MOdeling; Doms et al., 2011; Doms and Schättler, 2002; Baldauf et al., 2011) is a non-hydrostatic limited-area weather prediction model developed at the DWD. COSMO is used for operational, numerical weather predictions and research from the meso-β and to the meso-γ scales. COSMO is based on the primitive equations for compressible flow of moist air in the atmosphere, which are formulated in a rotated coordinate system and a vertical, generalized terrain-following grid. The rotation of the grid approximates a regular grid structure in order to avoid numerical instabilities. The primitive equations are translated into prognostic equations for wind (u $[m/s]$,v $[m/s]$,w $[m/s]$), temperature (T $[K]$), pressure deviation from a reference state (p' $[hPa]$), specific humidity (q_v $[kg/kg]$), specific cloud water and ice contents (q_l $[kg/kg]$, q_f $[kg/kg]$) and the specific water contents of rain and snow (q_r $[kg/kg]$, q_{sn} $[kg/kg]$). Diagnostic variables include, e.g., air density (ρ_a $[kg/m^3]$) and precipitation (P $[kg/m^2]$ or $[mm]$) (cf. Appendix F).

Numerically, these equations are implemented using the variable discretization of an Arakawa C-grid and a Lorenz vertical grid staggering. For the space discretization, a second-order finite differences scheme is used. Temporally, a modified Runge-Kutta time-splitting approach after Wicker and Skamarock (2002) is employed. Detailed descriptions of the

---

[1] http://tr32new.uni-koeln.de

dynamic equations and their numerical implementation in COSMO can be found in Doms et al. (2011).

Physical, subgrid-scale processes in COSMO are approximated using parameterizations. The most important parameterizations employed in COSMO are

- the simulation of subgrid-scale turbulence, which is based on a 2.5-level closure scheme accounting for condensation and thermal circulation effects;

- radiative transfer, which is simulated with a δ-two-stream approximation from Ritter and Geleyn (1992) using 3 shortwave and 5 longwave spectral intervals accounting for the full cloud-radiation feedbacks;

- moist convection, which is parameterized using the Tiedtke mass flux scheme (Tiedtke, 1989);

- and cloud microphysics, which are simulated with a bulk-water continuity model predicting cloud water, cloud ice, rain and snow; condensation and evaporation of cloud water are adjusted relative to saturation.

Boundary conditions and spectral nudging

COSMO is applied in a dynamical downscaling approach (e.g., Rockel, 2015), in which the information of global, large-scale models is transferred to smaller scales, using a higher-resolved, regional model domain. The higher resolution of the regional model allows to better resolve small scale variability of processes, land surface heterogeneity and topography, which can significantly influence the atmospheric stability, local weather and climate. The regional, limited-area domain of COSMO is nested into a global (re-)analysis or forecast, which provides the large-scale circulation as lateral boundary conditions via a Davis-type formulation (Davies, 1976). In this boundary value approach, a sponge zone around the regional domain is established (e.g., 8 grid cells in each direction), which forces the adjustment of the regional model state towards the outer, driving (re-)analysis or forecast. This adjustment is performed using a nudging coefficient γ $[1/s]$, which adjusts the interior state Ψ toward the outer state $\Psi*$ via $\gamma(\Psi*-\Psi)$, and decreases from the outer part of the sponge zone towards the regional domain (von Storch et al., 2000).

The forcing of the regional domain along the lateral boundaries is required for mathematical closure, but poses a number of difficulties (von Storch et al., 2000; Staniforth, 1997; Warner et al., 1997). The lateral forcing may constitute a significant loss of information, contained in the global simulations. This affects especially larger scales, which cannot be handled by a sponge zone of a few grid cells, and are the main contributor of error growth in a regional model (Vukicevic and Errico, 1990). Eventually, this leads to a distortion of the large-scale circulation in the interior, and prescribes a sensitivity of the regional model to domain size and position (Miguez-Macho et al., 2004). Therefore, studies proposed an additional forcing, the spectral nudging of the regional model domain as part of the dynamical downscaling. In this spectral nudging approach (Waldron et al., 1996; von Storch et al., 2000), large scales of the interior regional domain are adjusted towards the large scales of the global forcing, but allow an independent evolution of small scales by the regional model.

A brief description of the spectral nudging formulation, based on Waldron et al. (1996), von Storch et al. (2000) and Miguez-Macho et al. (2004), is given here.

The spectral nudging technique adds a term to the tendency equation of a prognostic variable Ψ, which nudges the original model solution $L(\Psi)$ towards the global solution at selected scales. The adjusted tendency equation is

$$\frac{d\Psi}{dt} = L(\Psi) + \sum_{n \leq N} \sum_{m \leq M} \alpha_{n,m}(A_{m,n} - A_{m,n}^o) \cdot exp(ik_m x) \cdot exp(ik_n y) \qquad (2.1)$$

with wave numbers n and m, and the spectral Fourier coefficients of the regional and the global model $A_{m,n}$ and $A_{m,n}^o$, respectively. The wave vector components $k_m = \frac{2\pi m}{D_x}$ and $k_n = \frac{2\pi n}{D_y}$ are defined by the domain size (D_x and D_y) and wave numbers m and n, which roughly represent the x- and y-direction. The nudging is performed on all wavenumbers smaller than N and M, with a nudging coefficient α. N and M are typically determined through the effective resolution of the global model. A nudging coefficient $\alpha_{n,m} = 0$ eliminates the spectral nudging effect, but maintains the one-way interaction at the lateral boundaries according to Davies (1976). A height-dependent nudging coefficient can be used, e.g., $\alpha(p < 850 \ hPa) = \alpha$ and $\alpha(p \geq 850 \ hPa) = 0$, and allows for a free evolution of boundary layer processes.

2.1.2 CLM

The NCAR Community Land Model Version 3.5 (CLM3.5; Oleson et al., 2008; Oleson et al., 2004) is the land surface scheme in TerrSysMP. CLM3.5 solves the energy balance at the land surface, providing the lower boundary condition for COSMO and the source/sink term of water for the surface-subsurface model ParFlow. The land surface is characterized by land cover types (vegetation, urban, wetland, lake, glacier). Subgrid heterogeneity in vegetation can be described using multiple plant functional types (PFTs; Bonan et al., 2002) per grid cell/soil column. Each PFT is defined by a number of plant specific physiological parameters (e.g., rooting depth and aerodynamic, optical, and photosynthesis properties), that are assumed constant in time. Additional characteristics of each PFT, such as the leaf area index (LAI), the stem area index (SAI), the monthly height at the top (MHTOP) and at the bottom (MHBOT), are prescribed on a monthly basis.

Based on these properties, CLM solves the surface energy equation for each PFT and grid cell,

$$R_{net} = H + LE + G \qquad (2.2)$$

which partitions the incoming net radiation (R_net [W/m^2]) into the sensible heat (H [W/m^2]), latent heat (LE [W/m^2]) and ground heat flux (G [W/m^2]). Note that this partitioning is strongly influenced by available soil moisture (Seneviratne et al., 2010), which highlights the connection to the deeper subsurface and water table depths. Here, the latent heat flux indicates the transfer of moisture from the land surface to the atmosphere, and can be converted to evapotranspiration (ET [kg/m^2] or [mm]) using the latent heat of

vaporization λ $[J/kg]$,

$$ET = \frac{1}{\lambda} \cdot LE = E_C + T_C + E_G \tag{2.3}$$

which subdivides into canopy evaporation (E_C $[kg/m^2]$), canopy transpiration (T_C $[kg/m^2]$) and ground evaporation (E_G $[kg/m^2]$).

CLM3.5 parameterizes root water uptake, interception, throughfall, canopy drip and snow accumulation and melt, which determine the sources and sinks of moisture supplied to the surface-subsurface model Parflow. The turbulent exchange between the land surface and the ambient atmosphere cannot be explicitly resolved and is approximated using the Monin-Obukhov (M-O) similarity theory (e.g., Foken, 2006), which is based on the observation of the logarithmic wind profile above a rough surface. The M-O theory limits the interaction to a 1D vertical exchange taking into account the mechanical generation of turbulence and the stability of the atmosphere.

2.1.3   ParFlow

The integrated surface-subsurface model ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006) simulates subsurface flow in variably saturated soil using the 3D Richards equation

$$S_S\theta(\psi_p)\frac{\partial \psi_p}{\partial t} + \phi\frac{\partial \theta(\psi_p)(\psi_p)}{\partial t} = \nabla q + q_S \tag{2.4}$$

with the specific storage coefficient $S_S[1/m]$, soil moisture θ [-], pressure head ψ_p $[m]$, porosity $\phi[-]$, an additional source/sink term q_S $[1/s]$, time t $[s]$ and the water flux q $[m/s]$. The Darcy water flux is described by

$$q = -k(x)k_r(\psi_p)\nabla(\psi_p - z) \tag{2.5}$$

with the saturated hydraulic conductivity $k(x)$[m/s], the relative permeability $k_r(\psi_p)$ [-] and the depth below the surface z $[m]$. The non-linear relationships of θ and k_r with ψ_p are described by the Van Genuchten model (Van Genuchten, 1980). The q_S term in the Richards equation is a general source/sink term that represents the exchanges with CLM (i.e., P and ET) and the inclusion of human water use (i.e., groundwater abstraction and irrigation).

In addition, ParFlow solves the 2D kinematic wave equation, which simulates overland flow and serves as an upper boundary condition for the Richards equation in the subsurface assuming pressure and flux continuity (Kollet and Maxwell, 2006). Overland flow velocities v_x $[m/s]$ and v_y $[m/s]$ in x- and y-direction are calculated as

$$v_x = \frac{\sqrt{S_{f,x}}}{n_m}\psi_s^{2/3}; \ v_y = \frac{\sqrt{S_{f,y}}}{n_m}\psi_s^{2/3} \tag{2.6}$$

and a respective outflow q_{out} $[m^2/s]$ at an outlet *outl* is defined as

$$q_{out} = \frac{\sqrt{S_{f,outl}}}{n_{m,outl}}\psi_s^{5/3} \tag{2.7}$$

with the ponding depth ψ_s $[m]$, the Manning's coefficient n_m $[s/m^{-1/3}]$ and the friction slopes $S_{f,x}$ [-], $S_{f,y}$ [-] and $S_{f,outl}$[-] in x- and y-direction and at the outlet, respectively.

A finite volume scheme in space, and an implicit backward Euler scheme in time are used to solve these equations in ParFlow. A terrain following grid is implemented and has been shown to improve simulations in case of coarse lateral resolution and large domains (Maxwell, 2013).

2.1.4 Coupling

The coupling between the three model components is performed using the Ocean Atmosphere Sea Ice Soil coupling toolkit (OASIS3-MCT; Valcke, 2013). The coupling in OASIS3-MCT and TerrSysMP is based on the MPMD paradigm, in which the individual models are executed separately and only interact at predefined coupling time steps through the exchange of fluxes and states. This modular setup allows the model components and executables to remain independent and enables the efficient use of parallel computing environments (Gasper et al., 2014). Practically, that means that COSMO, CLM and ParFlow are run in parallel and only interact through the exchange of fields through OASIS-MCT.

A brief description of the coupling between (i) COSMO and CLM, and (ii) CLM and Parflow is provided here and is based on Shrestha et al. (2014). Note that a new version of the coupling scheme has been introduced, which is described here. Figure 2.1 shows a vertical schematic of the variables exchanged between COSMO, CLM and ParFlow at every coupling time step.

COSMO passes the atmospheric state to the land surface model CLM (shortwave and longwave downward radiation (SW $[W/m^2]$ and LW_{DOWN} $[W/m^2]$), lowest-level temperature (T_a $[K]$), pressure (p $[hPa]$), specific humidity ($q_{v,a}$ $[kg/kg]$), wind speed ($|v_h|$ $[m/s]$), and precipitation (P $[mm]$)), while CLM returns the aerodynamic resistances of moisture, momentum and heat, surface humidity and temperature, and radiation components to COSMO (aerodynamic resistance of heat (r_H $[s/m]$), momentum (r_M $[s/m]$), and moisture (r_W $[s/m]$), surface temperature and humidity (T_s $[K]$, $q_{v,s}$ $[kg/kg]$), albedo (α_s [-]) and longwave, upward radiation (LW_{UP} $[W/m^2]$)). The aerodynamic resistances define the respective turbulent transfer coefficients,

$$C_H = \frac{1}{r_h|v_h|}; \ C_M = \frac{1}{r_m|v_h|}; \ C_W = \frac{1}{r_w|v_h|} \tag{2.8}$$

namely the coefficient of heat, the coefficient of momentum, and the coefficient of moisture, which in turn, allow an energy-, momentum and mass-conservative coupling. The horizontal wind velocity $|v_h| = (u^2 + v^2)$ $[m/s]$ is defined through the zonal and meridional wind components u $[m/s]$ and v $[m/s]$, respectively. The transfer coefficients are directly related to mass, energy and momentum, and define the latent heat flux,

$$LE = -C_W \rho_a \lambda |v_h|(q_{v,s} - q_{v,a}), \tag{2.9}$$

the sensible heat flux

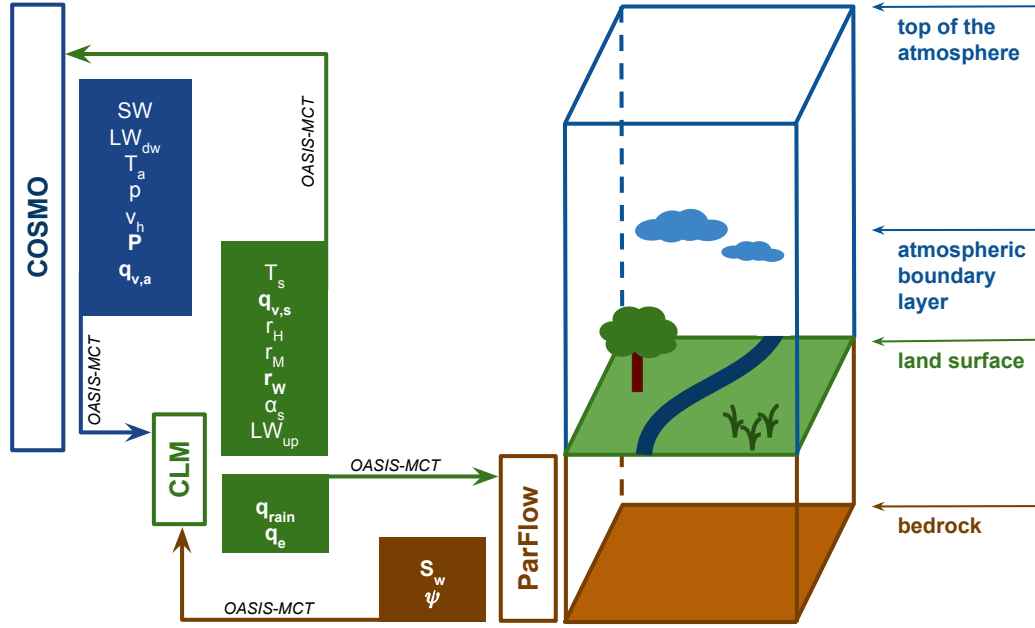$$H = -C_H \rho_a c_p |v_h|(T_s - T_a^*), \tag{2.10}$$

Figure 2.1: Schematic (not to scale) of the coupling in between the TerrSysMP model components COSMO, CLM and ParFlow using OASIS-MCT (redrawn after Shrestha et al., 2014). Variables that are directly related to the terrestrial water cycle are shown in bold.

and the zonal momentum flux $\tau [kg/m^2]$

$$\tau = C_M \rho_a u |v_h|, \tag{2.11}$$

based on the M-O theory. This includes the air density ρ_a $[kg/m^3]$, air and surface temperatures T_a^* $[K]$ and T_s $[K]$, air and surface humidity $q_{v,a}$ $[kg/kg]$ and $q_{v,s}$ $[kg/kg]$, the latent heat of vaporization λ $[J/kg]$ and the specific heat capacity for dry air c_p $[J/kgK]$. The air temperature is corrected for height differences between the lowest atmospheric level (T_a $[K]$) and the canopy,

$$T_a^* = \frac{T_a}{\Pi} \tag{2.12}$$

using the Exner function

$$\Pi = (p/p_0)^{(R_d/c_p)} \tag{2.13}$$

with pressure p $[hPa]$, reference pressure p_0 $[hPa]$ and the specific gas constant for dry air R_d $[J/kgK]$.

CLM and ParFlow are coupled through soil moisture in the upper 10 soil layers, which are setup identical in depth and porosity/soil texture. CLM provides the sources and sinks of moisture to ParFlow, namely the downward moisture flux from precipitation (q_{rain}), evaporation and transpiration by plants as depth differentiated sinks (q_e). ParFlow simulates 3D subsurface and overland flow, and provides the resulting pressure head (ψ) and profiles of relative saturation (S_w [-]) of the upper 10 layers to CLM. Through the latter, the parameterizations of runoff and 1D subsurface-drainage in CLM are replaced by the physics-based surface-subsurface flow from ParFlow.

Name	Res.	nlon x nlat	W	E	S	N
EUR11r	0.11°	424 x 412	-28.375	18.155	-23.375	21.835
EUR11i	0.11°	424 x 412	-44.812	65.187	21.812	72.687
EUR-11r-COS	0.11°	444 x 432	-29.475	19.255	-24.475	22.935
EUR11i-CLMPFL	0.11° (12.5 km)	436 x 424	-46.403	67.229	21.145	73.248

Table 2.1: Grid definitions of the European CORDEX domain, and the domain extensions used for COSMO, CLM and ParFlow. The index r refers to the rotated grid and the index i to the integrated grid.

2.2 TerrSysMP over Europe

The following paragraphs briefly describe the setup of TerrSysMP over Europe, and the workflow and computational resources required to run the fully-coupled modeling system.

2.2.1 The European CORDEX domain

TerrSysMP is setup over the European regional climate domain of the *Coordinated Regional Climate Downscaling Experiment* (CORDEX; Giorgi et al., 2009) project. The high-resolution implementation is chosen, which is 0.11° and represents a grid spacing of approximately 12.5 km. Table 2.1 gives an overview of the CORDEX specifications[2] for the integrated and rotated domain, respectively. The rotated grid with a north pole at 162°W and 39.25°N approximates a constant grid cell area. All model domains are slightly extended, i.e., from the CORDEX definition of 424 x 412 grid cells to 444 x 432 grid cells for COSMO, and to 436 x 424 grid cells for CLM and ParFlow (exemplary shown in Fig. 2.2). This extension allows to increase the sponge zone in COSMO, i.e., the transition between prescribed atmospheric states from the lateral boundary forcing and the dynamic model evolution, and minimizes the lateral boundary forcing effects for the inner domain. The 8 grid cells of the sponge zone were then removed from each side, and the inner 436 x 424 grid cells are coupled to CLM and ParFlow, which are setup with a constant grid cell area of 156 km^2 throughout the domain.

The domain covers Europe and extends towards Northern Africa in the South, and towards Georgia Republic and the Caspian Sea in the East (cf. Fig. 2.2). It consists to 47% of ocean and to 53% of land. Prominent European mountain ranges in this domain are the Alps (up to 3536 m in 0.11° resolution), the Carpathians, the Pyrenees, and the mountain ranges in Norway.

The vertical discretization of the models is constructed as follows. COSMO is setup with 50 vertical layers over the terrain following grid, reaching a height of approximately 22500 m (40 hPa). The lowest 10 levels cover approximately 1000 m. CLM3.5 has 10 soil layers of increasing thickness with depth, reaching a total depth of 3 m. These are shared with ParFlow, which is setup with 5 additional layers of increasing thickness below, reaching a total depth of 103 m.
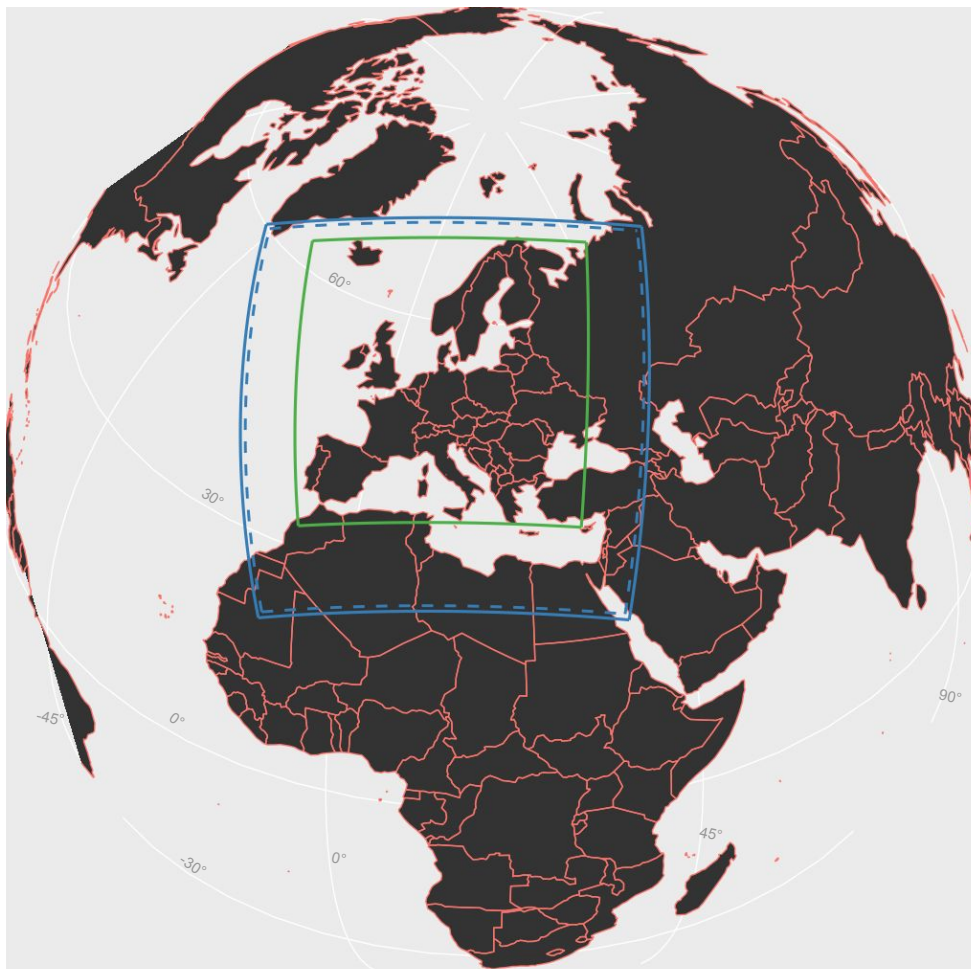
Figure 2.2: Projection of the European CORDEX domain on the globe. Lines show the COSMO domain (solid blue line), the ParFlow-CLM domain (dashed blue line) and the focus domain for analysis (solid green line).

2.2.2   External input data sets

A large number of external data sets is required to implement a model domain with Terr-SysMP. An overview of the data sets used for the setup over the European CORDEX domain is given in Table 2.2, subdivided into the compartments, and briefly described here.

The topography in ParFlow is represented by D4 slopes calculated from the digital elevation model. The subsurface characteristics are described by the soil class distribution from the Food and Agriculture Organization (FAO) and a deeper subsurface data set from Gleeson et al. (2011a, 2011b), which accounts for consolidated and unconsolidated geology below the upper soil layers. Both data sets are used to setup two hydrofacies distributions (HFDs) and study the impact of a vertically heterogeneous soil profiles on land-atmosphere feedbacks. Soil parameters, such as saturated hydraulic permeability, porosity, and Van Genuchten parameters, are obtained using pedotransfer functions from Schaap and Leij (1998). A detailed description of the employed HFDs can be found in Chapter 3. In Chapter 4 and 5, human water use data sets from Wada et al. (2012), Wada et al. (2016a), and Siebert and Döll (2010) and Siebert et al. (2010) are employed. These comprise realistic

---

[2]http://cordex.dmi.dk/joomla/images/CORDEX/cordexarchivespecifications.pdf

estimates of irrigation and groundwater abstraction for irrigation, domestic and industrial water use. Two data sets are used to study the sensitivity of land-atmosphere feedbacks to human water use estimates.

Vegetation in CLM is described by plant functional types (PFTs) from the Moderate Resolution Imaging Spectroradiometer (MODIS, Friedl et al. 2002), and corresponds to land cover/land use. To account for subgrid-scale heterogeneity at the 156 km^2 pixel, a percentage of the 4 most dominant PFTs per grid cell is used. In a first setup (simulations from Chapter 3), the plant physiological parameters for each PFT were derived from the global CLM surface data set (Oleson et al., 2008). The European mean of each parameter (MHTOP, MHBOT, LAI, SAI) for each PFT was used to describe seasonality and is constant over the domain. For the simulations in Chapter 4 and 5, the LAI was updated using the MODIS LAI estimates for the year 2003, providing a spatially varying physiologic description for the same PFT.

2.2.3 Time steps and coupling

COSMO is run with a 60 s time step for all simulations, and the radiation scheme is called every 15 min. In Chapter 3, ParFlow and CLM are run with time steps and a coupling frequency of 1 h. For the simulations in Chapter 4 and 5, the time steps of CLM and ParFlow and the coupling frequency are decreased to 180 s in order to to better resolve non-linearities in land-atmosphere feedbacks at the sub-hourly time scale.

All models are constructed using the same spatial resolution, thus, no interpolation or downscaling between the individual model components is required. CLM and ParFlow are setup on a regular grid and coupling with COSMO is performed using the rotated grid with a regular grid spacing. A novelty compared to the previously used model domains, such as the Rur catchment (e.g., Shrestha et al., 2014; Simmer et al., 2015; Rahman et al., 2015; Sulis et al., 2017), is the incorporation of a land/sea-mask in the coupling. Exchange with the ocean is simulated with COSMO, using prescribed sea surface temperatures from the driving lateral boundary conditions.

2.2.4 Initial and boundary conditions

Two assumptions are applied for the setup of ParFlow over the European CORDEX domain. First, it is assumed that there is no outflow at the bottom of the subsurface model, which mimics the behavior of impermeable bedrock at a depth of 103 m in all simulations. Here, a no-flow lower boundary condition is applied. Along the coastlines and the continental shelves, lateral outflow is allowed through the use of a Dirichlet lateral boundary condition, which mimics the sea surface through a hydrostatic profile with the water table at shallow depth below the surface. Especially the latter constitutes a major simplification, because the exchange of groundwater at the ocean-freshwater interface is characterized by complex processes related to variable density flow and local heterogeneity, which cannot be resolved in the applied subsurface model.

---

[3](https://lta.cr.usgs.gov/GTOPO30)

Name	Variables	References
Surface/Subsurface		
GTOPO30	topography topographic slopes	USGS[3]
FAO soil database	hydraulic conductivity porosity %sand %clay	Food and Agriculture Organization (1988)
Gleeson database	hydraulic conductivity	Gleeson et al. (2011a) Gleeson et al. (2011b)
Water table depth	WTD	Fan et al. (2013)
Human water use	groundwater abstraction irrigation	Wada et al. (2012) Wada et al. (2016a) Siebert and Döll (2010) Siebert et al. (2010)
Vegetation		
global CLM surface data set	stem area index leaf area index monthly height top monthly height bottom	Oleson et al. (2008)
MODIS	land use/ PFTs leaf area index	Friedl et al. (2002)
Atmosphere		
ERA-Interim	specific humidity horizontal wind vertical wind temperature cloud liquid water cloud ice content geopotential sea surface temperature	Dee et al. (2011)

Table 2.2: Overview of the external data sets used for the setup of TerrSysMP over the European CORDEX domain.

The setup of TerrSysMP requires an initial condition of the subsurface state, which is highly uncertain due to missing observations of, e.g., water table depths (WTD) and soil texture heterogeneity at different depths and spatial resolutions. In order to overcome this issue, a multi-year spinup of the coupled surface-subsurface model (ParFlow-CLM) was performed, as initial profiles of, e.g., pressure head at each grid cell are required. A spinup describes the process of re-running a numerical model over a chosen time period from an (arbitrary) initial condition until an equilibrium between the subsurface states and fluxes with the land surface and the atmospheric forcing is reached. The spinup process eliminates noise and spurious signals due to inconsistencies between the preliminary initial state, the model setup and the atmospheric forcing. The ParFlow-CLM spinup started from a global water table depth data set by Fan et al., (2013), which was compiled using well observations and a simplified groundwater model. As the model setup between this global groundwater model and TerrSysMP diverges, the multi-year spinup with atmospheric forcing from hindcast simulations over the same domain, from Vautard et al. (2013) and Kotlarski et al. (2014), was performed. The spinup over the year 2003 was repeated until a dynamic equilibrium of the surface-subsurface state, which is physically consistent with the land surface processes and the atmospheric forcing, was achieved. Setups with different HFDs were spun-up separately. Simulations with CLM and ParFlow are started using these initial conditions.

For the atmospheric compartment, COSMO is initialized with ERA-Interim reanalysis (Dee et al., 2011) at the start of the simulation. 3-hourly ERA-Interim data sets are used as lateral boundary conditions. Therefore, additional 3, 6 and 9h forecasts from ERA-Interim analysis at 00 and 12 UTC are used. In Chapter 3, the atmosphere is re-initialized every 24 hours at 00 UTC. In Chapter 4 and 5, the simulations are only initialized once, and additionally employ spectral nudging (von Storch et al., 2000; Miguez-Macho et al., 2004), which keeps the large-scale circulation between the simulations scenarios consistent to the driving ERA-Interim reanalysis. This minimizes the differences, arising through different synoptic situations over the continental-scale domain. The spectral nudging is applied to horizontal wind components (u,v) above the planetary boundary layer ($p < 850hPa$) every $300s$ with a nudging coefficient $\alpha = 0.05$ for wave numbers smaller than 14 (approximately $200km$). The wave number is determined from the effective resolution of the driving ERA-Interim forcing, and an evaluation of the kinetic energy spectra from experiments with varying wave numbers.

In this thesis, two versions of COSMO are used: the simulations in Chapter 3 are performed using COSMO Version 4.21 (TerrSysMP Version 1.1.0MCT), while the simulations in Chapters 4 and 5 are performed using COSMO Version 5.1 (TerrSysMP Version 1.2.0MCT). The model versions are switched due to the better compute time performance of the spectral nudging technique in COSMO Version 5.1.

2.2.5 Supercomputing environments and workflow

Two high-performance computing (HPC) facilities are used to perform the simulations in this thesis. The main HPC environments are (i) the supercomputer JUQUEEN at the Juelich
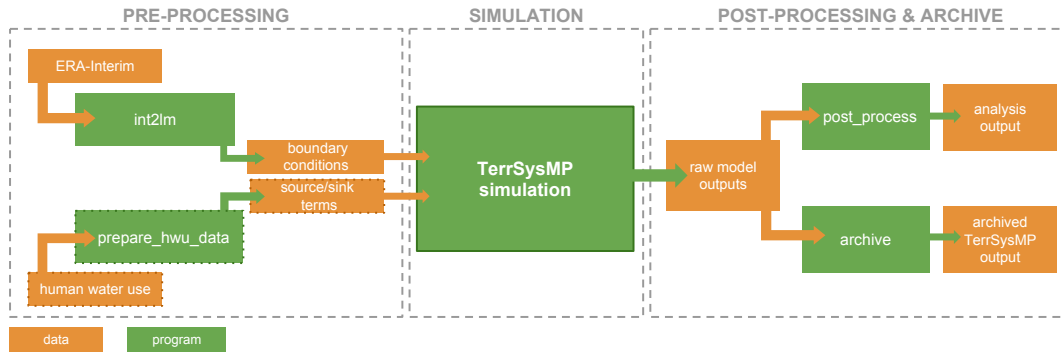
Figure 2.3: Schematic of the applied simulation workflow in this thesis. TerrSysMP simulations consider the fully-coupled system, i.e., ParFlow-CLM-COSMO. Orange boxes indicate input/output data; green boxes indicate programs and routines.

Supercomputing Centre (JSC[4]) and (ii) the supercomputer CRAY XC-40 at the European Centre for Medium-range Weather Forecasts (ECMWF[5]). The simulations in Chapter 3 are performed on JUQUEEN, while the simulations in Chapter 4 and 5 are performed on the CRAY at ECMWF.

The simulation workflow consists of three main steps, namely pre-processing, followed by the actual TerrSysMP simulation, and a final post-processing and archiving procedure. The workflow is illustrated in Fig. 2.3. The pre-processing includes the retrieval of ERA-Interim reanalysis data and interpolation with *int2lm* (Schättler, 2013) to boundary conditions for COSMO. For the simulations in Chapter 4 and 5, this also encompasses the retrieval and preparation of the human water use data sets into a time-consistent and ParFlow-readable data format. Using these data sets and the initial conditions for ParFlow and CLM, the TerrSysMP simulations are performed. The output contains (i) hourly COSMO files, (ii) hourly ParFlow files for saturation and pressure head, and (iii) daily CLM files. In a first step, these are merged to monthly files for a limited number of variables of interest. The post-processing also includes the calculation of, e.g., water table depths from the ParFlow output. These files are used in the ensuing analysis. In a second step, the raw model outputs are archived. During the course of this thesis, this workflow has been implemented in ECMWF's workflow package ecflow[6], which also includes developed routines to restart the simulations enabling long-term simulations beyond the maximum wallclock time of supercomputers.

2.2.6   Computational resources and storage requirements

The setup and performance of the simulations with TerrSysMP is computationally demanding. In addition to the simulations presented in this thesis, the setup, the spinup and the data handling require extensive computational resources. The following paragraphs provide a short overview of average simulation statistics.

In this thesis, the model setup and simulations used in total more than 25 Mio core-hours on the HPC facilities presented above. Table 2.3 gives an overview of the average simulation

---

[4]http://www.fz-juelich.de/ias/jsc/EN/Expertise/Supercomputers/JUQUEEN/JUQUEEN_node.html
[5]https://www.ecmwf.int/en/computing/our-facilities/supercomputer
[6]https://confluence.ecmwf.int/display/ECFLOW/ecflow+home

Simulation	Nodes (#processors)	HWU	wallclock time	SBU (core-hours)
TerrSysMP	14 (504)	no	155h	1.223.067 (75.919)
TerrSysMP-HWU	14 (504)	yes	188h	1.530.401 (94.997)

Table 2.3: Average runtimes and core-hours for a one-year simulations with TerrSysMP on 14 nodes with 36 processors per node (504 processors: 144 processors for ParFlow, 324 processors for COSMO and 36 processors for CLM), with and without the incorporation of human water use, on the CRAY XC-40 using spectral nudging (m and $n = 14$, $\alpha = 0.05$, time step=300 s) and a coupling frequency of 180 s.

times and resources for a one-year simulation with TerrSysMP over the European CORDEX domain at 0.11° spatial resolution. Note that this is only an average estimate for a specific setup on the CRAY at ECMWF, and that single simulation times vary depending on, e.g., the specific HPC resources used, the coupling frequency and the time step of the spectral nudging, and the number of iterations required to solve the non-linear system of equations in ParFlow. The latter is strongly dependent on, e.g., spatial heterogeneity, water stress and overland flow.

An average simulation, as presented in Chapter 4 and 5, required between 75.000 and 100.000 core-hours, which illustrates the need of massively parallel HPC environments to perform this type of study. In addition, the compute time required for (i) the multi-year spinup, (ii) pre- and post-processing of large data sets, and (iii) testing must be considered.

Model inputs for COSMO is 61 GB per year for ERA-Interim data, and 368 GB per year for the interpolated boundary conditions. A single human water use data set with daily estimates is about 3 GB per year in the raw format, but accumulates to approximately 1 TB per year in the pre-processed format for TerrSysMP, which requires a file for each time step. The raw model output for a one-year simulation with TerrSysMP over the European CORDEX domain reaches 8.4 TB per year, with 7.2 TB for the 3D and 2D fields of COSMO, 0.8 TB for the 3D and 2D fields of CLM, and 0.4 TB for the 3D pressure and saturation fields of ParFlow. A list of model output variables that are archived can be found in Appendix F.

Chapter 3

Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heatwave in 2003

Abstract

The impact of 3D groundwater dynamics as part of the hydrologic cycle is rarely considered in regional climate simulation experiments. However, there exists a spatial and temporal connection between groundwater and soil moisture near the land surface, which can influence the land surface-atmosphere feedbacks during heatwaves. This study assesses the sensitivity of bedrock-to-atmosphere simulations to groundwater representations at the continental scale during the European heatwave 2003 using an integrated fully-coupled soil-vegetation-atmosphere model. The analysis is based on the comparison of two groundwater configurations: (1) 3D physics-based variably saturated groundwater dynamics and (2) a 1D free drainage (FD) approach. Furthermore, two different subsurface hydrofacies distributions (HFD) account for the uncertainty of the subsurface hydraulic characteristics, and ensemble simulations address the uncertainty arising from different surface-subsurface initial conditions. The results show that the groundwater representation significantly impacts land surface-atmosphere processes. Differences between the two groundwater configurations follow subsurface patterns and the largest differences are observed for shallow water table depths. While the physics-based setup is less sensitive to the HFD, the parameterized FD simulations are highly sensitive to the hydraulic characteristics of the subsurface. An analysis of variance shows that both, the groundwater configuration and the HFD, induce variability across all compartments with decreasing impact from the subsurface to the atmosphere, while the initial condition has only a minor impact.

Key points

- Continental modeling of the water and energy cycle from groundwater into the atmosphere

- Impact study of groundwater dynamics and parameterizations on atmospheric processes

- Land surface and atmospheric variability is induced by the groundwater configuration

3.1 Introduction

The role of soil moisture in climate, and, particularly, the effect of soil moisture on atmospheric processes has been subject of many studies (Seneviratne et al., 2010; Betts et al., 1996), based on observations (Hirschi et al., 2010; Miralles et al., 2012; Guillod et al., 2014) and explored with the help of numerical experiments (Jaeger and Seneviratne, 2010; Vautard et al., 2007). Moreover, land-atmosphere interactions (Seneviratne et al., 2006; Dirmeyer, 2011), such as the soil moisture-temperature and the soil moisture-precipitation feedback, are assumed to further intensify the hydrologic cycle (Huntington, 2006) along with dry and wet spells (Mueller and Seneviratne, 2012).

Yet, the explicit interconnection between groundwater, the land surface and the atmosphere is generally neglected or highly simplified, and the hydrologic formulations used in regional climate simulation experiments are typically based on one-dimensional parameterizations in the vertical direction. In these parameterizations, a free-drainage approach (e.g., Campoy et al., 2013) is commonly used as a bottom boundary condition and ponded water at the land surface is removed. These simplified parameterizations are a source of uncertainty in the simulation of soil moisture and involve fundamental restrictions, e.g., saturation can only occur if the rainfall rate exceeds the infiltration capacity, which leads to overestimated drainage especially in wet seasons (Fan, 2015).

However, a common hypothesis is that groundwater processes significantly influence land- atmosphere interactions, atmospheric processes and patterns in water and energy fluxes and states across different temporal scales (from weather to climate) and spatial scales from catchments to the continent (e.g., Taylor et al., 2013). Here, the following rationale is applied: There exists a hydraulic connection of the deeper subsurface to the vadose zone beyond the extents of the catchment scale (e.g., Schaller and Fan, 2009; Krakauer et al., 2014; Fan, 2015), which further significantly influences the land surface through surface-subsurface interactions. Finally, this connection affects large-scale patterns, processes and extremes in the atmosphere through land surface-atmosphere feedbacks. Consequently, there is a need to either further improve the representation of groundwater in land surface models (LSMs) (Clark et al., 2015; Hazenberg et al., 2016) or to incorporate fully integrated, physics-based groundwater models in LSMs and large-scale climate simulations. As the representation of hydrologic processes in land surface models is under ongoing improvements, the focus in this study is placed on the integration of physics-based groundwater models into Earth System Models (ESMs) (Giorgi, 1995). Recent technological developments afford to overcome the computational burden associated with large-scale hydrologic modeling (Bierkens et al., 2015; Bierkens, 2015) and the integration of sophisticated numerical groundwater models into continental to global ESMs. In this context, high-resolution physics-based hydrologic models are being developed and used to examine land-energy interactions at the continental scale (Maxwell et al., 2015; Condon and Maxwell, 2015) but do not consider the aforementioned atmospheric feedbacks yet.

However, there already exist a number of modeling studies, which simulate the full hydrologic cycle from the bedrock into the atmosphere including two-way feedbacks between the different compartments. These studies started from "idealized" and simplistic

atmospheric conditions (York et al., 2002) or models considering parameterized groundwater approaches (e.g., Walko et al., 2000; Seuffert et al., 2002; Miguez-Macho and Fan, 2012) up to fully-coupled soil-vegetation-atmosphere systems including improved groundwater approaches also at the continental scale (Anyah et al., 2008). Nevertheless, numerical groundwater models solving the full 3D Richards equation are rarely used and most of these modeling studies are restricted to the catchment scale (e.g., Maxwell et al., 2007, 2011; Shrestha et al., 2014; Butts et al., 2014; Larsen et al., 2014; Rahman et al., 2015; Davison et al., 2015).

Recently, Barlage et al. (2015) presented a study, which explores the effect of different land surface models and groundwater parameterizations on six months regional climate simulations with the WRF model over the continental U.S. This study indicates that the incorporation of groundwater processes at the continental scale, even at a 30 km resolution, can alleviate a warm/dry bias up to $1 - 2\ °C$. Especially heatwaves, such as the European heatwave in 2003 (Black et al., 2004; Schär and Jendritzky, 2004), have been subject of many studies concerning predictability with focus on land-atmosphere feedbacks (Zaitchik et al., 2006; Della-Marta et al., 2007; Vautard et al., 2007; Jaeger and Seneviratne, 2010; Weisheimer et al., 2011; Stéfanon et al., 2013). The European heatwave 2003 exhibited exceptionally high temperature anomalies with record-breaking heat days in the first half of August (Black et al., 2004). The causes for this heatwave are controversially discussed (Weisheimer et al., 2011), but most studies indicate that a precipitation deficit lead to dry soil moisture anomalies from May until August 2003, significantly enhanced land-atmosphere feedbacks and contributed to the extreme summer 2003. Particularly, MacLeod et al. (2016) showed that under such extremely dry conditions, the main uncertainty in land surface models arises from the hydraulic characteristics of the soil; and seasonal predictions in strongly coupled regions can be significantly improved by incorporating this uncertainty in the hydrology parameterizations of the respective land surface models. Additional numerical experiments and analyses over the heatwave 2003 have been conducted to account for different soil moisture initializations (Fischer et al., 2007; Ferranti and Viterbo, 2006) and anomalously dry soil conditions (Fennessy and Kinter, 2011). Yet, these studies employed land surface models with parameterized groundwater processes, used artificial perturbations of soil moisture initializations, and neglected the aforementioned subsurface-surface connection. The question arises whether the 3D hydrological processes in the model systematically affect land-atmosphere feedbacks in addition to the soil moisture initialization. In other words, does the groundwater representation affect atmospheric variability including extremes, such as the heatwave in 2003, through the hydrologic cycle and through feedback mechanisms at the continental scale? The following hypotheses are formulated: First, the groundwater representation significantly influences land surface and atmospheric processes during the heatwave 2003. In other words, there exists an additional uncertainty in regional climate simulations due to the groundwater representation. Secondly, groundwater dynamics alleviate the evolution of extremes through attenuated land-atmosphere feedbacks during this heatwave.

This study attempts to test these hypotheses with a sensitivity study using an integrated fully-coupled soil-vegetation-atmosphere model over the European continent during

one month of the European heatwave 2003. The Terrestrial Systems Modeling Platform, TerrSysMP (Shrestha et al., 2014), with two different groundwater configurations,

1. physics-based, 3D variably saturated groundwater dynamics, and

2. a 1D free drainage approach,

is set up over the European CORDEX domain (Fig. 3.1). Ensemble simulations are conducted utilizing different initial surface-subsurface states. Two different subsurface hydrofacies distributions (HFD) are used to account for the uncertainty of the subsurface hydraulic characteristics. The differences between the fully-coupled simulations, which differ only through the groundwater formulation or the HFD, are analyzed and the uncertainty of the resulting feedback signals is exploited using the analysis of variance method. Additionally, the resulting fields are compared to atmospheric and land surface reanalyses, such as ERA-Interim (Dee et al., 2011) and ERA-Interim/Land (Balsamo et al., 2015).

A brief description of TerrSysMP and its setup over the European CORDEX domain is provided in section 3.2. The experiment design is described in section 3.3. Section 3.4 illustrates and discusses the uncertainty of the performed simulations and utilizes the subsurface-land surface-atmosphere feedbacks with respect to groundwater dynamics. Differences to ERA-Interim and ERA-Interim/Land are presented. Finally, a summary and conclusions are provided in section 3.5.

3.2 Modeling system

3.2.1 TerrSysMP

TerrSysMP (Shrestha et al., 2014, Gasper et al., 2014) consists of the numerical weather prediction model COSMO (COnsortium for Small Scale MOdelling, Version 4.21) (Baldauf et al., 2011; Doms and Schättler, 2002) from the German Weather Service, the land surface model CLM (National Centre for Atmospheric Research Community Land Model, Version 3.5) (Oleson et al., 2008) and the surface-subsurface model ParFlow (Parallel Flow, Version 3.1) (Kollet and Maxwell, 2006; Jones and Woodward, 2001; Maxwell, 2013), which are coupled via the Ocean Atmosphere Sea Ice Soil coupling tool OASIS3-MCT (Valcke, 2013). The non-hydrostatic limited-area numerical weather prediction (NWP) model COSMO, simulates the atmospheric processes within TerrSysMP. COSMO sends the short- and long-wave radiation, as well as near-surface temperature, pressure, specific humidity, wind and precipitation to CLM3.5. The one-dimensional land surface model CLM3.5 serves as the lower boundary for the atmospheric model COSMO simulating the land surface energy balance. It also provides the sources and sinks of soil moisture for the surface and subsurface model ParFlow, a three-dimensional variably saturated surface-subsurface flow code, which simulates close-to-physics, 3D groundwater dynamics.

3.2.2 Setup of TerrSysMP over the European CORDEX domain

TerrSysMP is set up over the European continent illustrated in Fig. 3.1. To ensure comparability with a large community, the grid definitions from the COordinated Regional
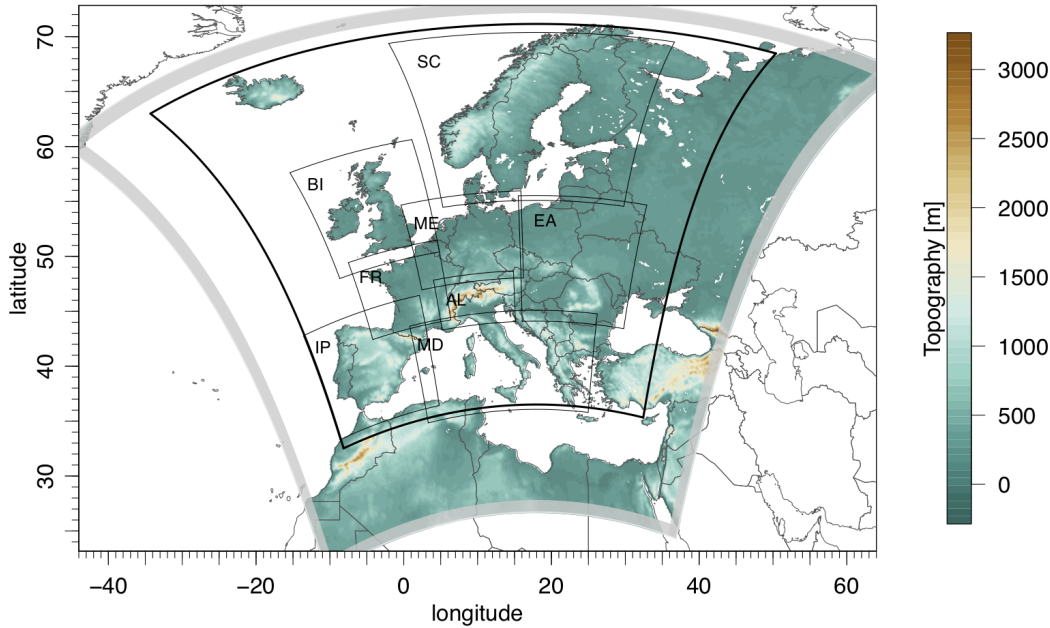
Figure 3.1: Topography [m a.m.s.l.] over the European CORDEX domain at 0.11° resolution. The gray boundary illustrates the boundary relaxation zone used and the black box in the inner domain indicates the focus domain for analysis. The small inner boxes show the PRUDENCE regions and the respective abbreviation.

Downscaling EXperiment (CORDEX) framework (Giorgi et al., 2009) are applied. The respective high-resolution grid has a horizontal grid size of 0.11° (12 km, EUR-11). Simulations are performed over the entire CORDEX domain, while analyses are performed on a focus domain, exemplarily depicted in Fig. 3.1. Moreover, we make use of pre-defined regions from the 'Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects' (PRUDENCE, Christensen et al., 2007) project for comprehensive analyses. We refer to these regions as the PRUDENCE regions.

COSMO

In this setup, COSMO is nested within ERA-Interim (Dee et al., 2011), the reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA-Interim provides the initial state and the lateral boundaries of the atmosphere to COSMO. The ERA-Interim 3D fields (specific humidity, temperature, wind, cloud liquid and ice water content) in 0.5° lateral resolution as well as the geopotential are downscaled to the CORDEX-EUR11 grid of TerrSysMP and serve as initial conditions and lateral boundaries. In order to update the lateral boundaries more frequently than ERA-Interim analyses are available, additional 3, 6 and 9 h forecasts from ERA-Interim analysis at 00 and 12 UTC are used, respectively.

The following physical parameterizations are applied to account for atmospheric subgrid-scale processes. Convection is parameterized via the Tiedtke mass flux scheme (Tiedtke, 1989). For the radiative transfer, the δ-two-stream approximation of the radiative transfer equation according to Ritter and Geleyn (1992) is used. Vertical turbulent diffusion is simulated with a 2.5-level closure scheme based on the prognostic equation for turbulent

	ParFlow	CLM3.5	COSMO
Horizontal extensions	436 x 424	436 x 424	444 x 432
# of active grid cells	99.395	99.395	191.808
Horizontal resolution	0.11°	0.11°	0.11° (rot.)
Vertical layers	15	10	50
Vertical resolution	variable	variable	variable
Depth/height	103 m	3 m	21 km
Time step size	3600 s	3600 s	60 s

Table 3.1: Model setup of TerrSysMP over the European CORDEX domain in 0.11° resolution.

kinetic energy and cloud micro-physics are simulated with a bulk-water continuity model, predicting cloud water, cloud ice, rain and snow.

CLM

The land surface composition for CLM3.5 over the European CORDEX domain is based on the Moderate Resolution Imaging Spectroradiometer (MODIS) data set (Friedl et al., 2002), and the land use is transferred to plant functional types (PFT). Due to the coarse resolution of the land surface, the setup of TerrSysMP accounts for sub-grid heterogeneity by employing multiple land units for each grid cell (through the tiling approach in CLM3.5). Additional parameters for each PFT, such as the leaf area index (LAI), the stem area index (SAI) and the monthly bottom and top heights of each PFT (MHBOT and MHTOP), are calculated based on the global CLM3.5 surface data set (Oleson et al., 2008). The soil texture (i.e., percent of sand and clay in the soil) is defined according to the soil properties in ParFlow.

ParFlow

ParFlow is setup with 15 vertical layers, resulting in a total depth of 103 m with increasing thickness towards the bottom. The thickness of the upper 10 layers is identical to the layers in CLM3.5, reaching a depth of 3 m. A Dirichlet lateral boundary condition with a hydrostatic profile and the water table at the top is set for the hydraulic pressure along the coastlines. The topography in ParFlow is represented by D4 slopes calculated from the digital elevation model (DEM) and the terrain following grid transform with variable vertical discretization (Maxwell, 2013) is used in order to improve the simulations for large topographic gradients and coarse lateral resolutions. The setup of ParFlow, i.e., the ground-water representation and hydraulic characteristics, is varied in this study and described in detail in section3.3.

Coupling

The time step for ParFlow and CLM3.5 is hourly, while COSMO runs with a 60 s time step. Coupling between the component models is applied at an hourly frequency with averaged

values from COSMO. Table 3.1 summarizes the computational setup of each model over the CORDEX domain.

3.3   Experiment design

This sensitivity study consists of two different groundwater representations (Fig. 3.2) and two HFD within ParFlow, and an ensemble of varying subsurface initial conditions resulting in a total number of 20 simulations.

3.3.1   Groundwater configuration

Two groundwater configurations are setup over the European CORDEX domain. In the first configuration, a physics-based groundwater configuration is applied. Within this setup, which is called TerrSysMP(3D), ParFlow solves the 3D-Richards' equation which describes the variably saturated groundwater flow (Kollet and Maxwell, 2006):

$$S_S\theta\frac{\partial\psi_p}{\partial t} + \phi\frac{\partial\theta(\psi_p)}{\partial t} = \nabla q + q_S \tag{3.1}$$

with the water flux q $[m/s]$ as $q = -k(x)k_r(\psi)\nabla(\psi - z)$, soil moisture θ [-] and ψ $[m]$ as the subsurface pressure head, the depth below the surface z $[m]$, the specific storage coefficient S_S $[1/m]$, porosity ϕ [-], saturated hydraulic conductivity $k(x)$ $[m/s]$ and k_r [-] as the relative permeability. q_S $[1/s]$ represents a source/sink term of water fluxes arising from precipitation and vegetation processes, such as evapotranspiration, simulated by CLM3.5. Surface and subsurface are coupled through the boundary condition of the shallow overland flow, which is described by the 2D-kinematic wave equation (Kollet and Maxwell, 2006).

In the second configuration, the complexity of ParFlow is reduced and a vertical, one-dimensional "free drainage" or "(open) bucket" model similar to Niu et al. (2007) and Campoy et al. (2013), is applied. In this case, no lateral subsurface flow is simulated and only the 1D-Richards' equation

$$\phi\frac{\partial\theta(\psi_p)}{\partial t} = \frac{\partial q}{\partial z} \tag{3.2}$$

is solved. Additionally, the lower boundary condition is changed and a one-dimensional gravitational soil water flux q_g $[m/s]$ is applied at the lower bottom with a rate defined by the hydraulic conductivity $k(x)$ and the relative permeability k_r,

$$q_g = -k(x)k_r(\psi)\frac{\partial\psi_g}{\partial z} = k(x)k_r(\psi)\frac{\partial z}{\partial z} = k(x)k_r(\psi) \tag{3.3}$$

with ψ_g $[m]$ as the gravitational potential which is equal to the depth z. This groundwater configuration within the fully-coupled system is called TerrSysMP(FD), where FD represents the gravitational free drainage. In this 1D approach, no horizontal movement of water is considered, and capillary upward fluxes from a water table, are neglected. Consequently, saturation can only occur if the precipitation rate is equal to or exceeds the infiltration rate (Fan, 2015).
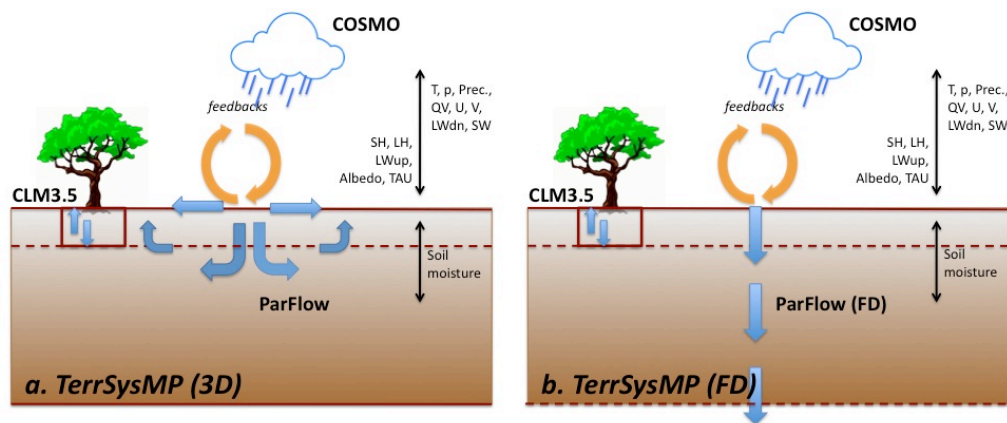
Figure 3.2: Sketch of the two groundwater configurations implemented in TerrSysMP and used in this study: (a) the complex, 3D and physics-based setup of ParFlow within the fully-coupled system, (b) a 1D free-drainage approach of ParFlow within the fully-coupled system.

3.3.2 Hydro-facies distributions

In addition to the groundwater configuration, the subsurface hydraulic characteristics are varied by applying two HFDs:

- HFD1: the entire soil column is described by the Food and Agriculture Organization (FAO) soil database (FAO, 1988)

- HFD2: the FAO soil database is used for the upper soil ($3\,m$) and the Gleeson database (Gleeson et al., 2011a,b) for the deeper subsurface representing the shallow aquifers and the bedrock.

While the first ten layers are identical in HFD1 and HFD2 and consequently in the setup of CLM3.5, the deeper subsurface in HFD2 exhibits mostly decreased hydraulic conductivities and is subsequently less permeable and less conductive compared to HFD1, except for some highly conductive parts, e.g., in Mid-Europe. In all cases, the respective fields are interpolated to the EUR11 grid and the pedo-transfer functions from Schaap and Leij (1998) are used to obtain the respective parameters, i.e., saturated hydraulic permeability, porosity and van Genuchten parameters. Figure 3.3 shows the hydraulic conductivity $k(x)$ for both databases, FAO and Gleeson. The patterns from the Gleeson database specifically for the deeper subsurface follow the geologic features of Europe, e.g., the highly conductive part from Northern Mid-Europe to Belarus was formed during the Holocene and the glacial drift.

3.3.3 Initial condition

In order to account for the uncertainty of the subsurface initial state, an ensemble is constructed by perturbing the subsurface initial conditions (i.e. saturation) as follows: First, spinup simulations of the surface and subsurface are performed in order to achieve a dynamic equilibrium concerning groundwater dynamics in TerrSysMP(3D). The equilibrium water table depth proposed by Fan et al. (2013) is used to start multi-year simulations of
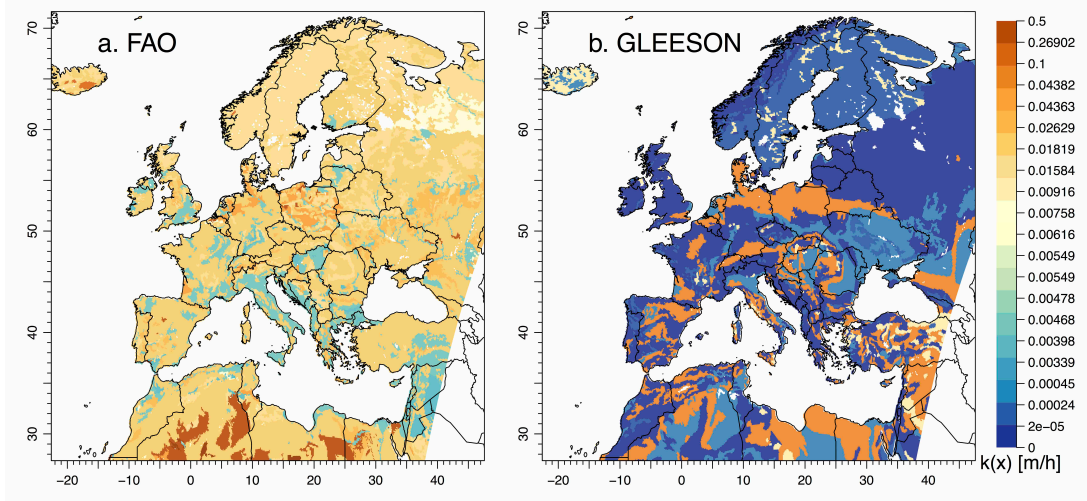
Figure 3.3: Spatial distributions of the hydraulic conductivity $k(x)$ $[m/h]$ prescribed for (a) the FAO and (b) the Gleeson database.

the surface and subsurface instead of using arbitrary initial conditions. For initialization, a hydrostatic profile around this equilibrium water table depth is assumed and repeated multi-year offline simulations with atmospheric forcing for the year 2003 were performed in order to adapt the hydrological compartments of TerrSysMP to the extreme dry hydro-meteorological conditions in 2003 and to account for the long-term soil moisture memory. Samples of the respective states of this simulation then serve as the initial condition for the land surface and the subsurface in the fully-coupled simulations. Five different states within the dry period and around the start of the simulation period (14th of July, 23rd of July, 1st of August, 10th of August and 15th of August) are chosen in order to account for the uncertainty of the subsurface initial conditions and to expand the sample size of the subsequent analyses. Note that the long-term spinup has been performed with the subsurface defined by HFD1. Subsequently, an additional, shorter spinup with HFD2 has been performed and the respective samples from this spinup are used to initialize fully-coupled simulations with subsurface characteristics from HFD2 consistently.

3.3.4   Simulation period and flow chart

Figure 3.4 illustrates the daily $2\,m$ temperature from ERA-Interim exemplarily averaged over the PRUDENCE region France in 2003. While large parts of Europe were affected by this heatwave, Mid-Europe and France achieved the highest temperatures at the beginning of August, lasting for 2 weeks. Here, the largest impact of the groundwater representation is expected. In order to examine the influence of the subsequent precipitation events on the respective ensemble spread and the difference between the groundwater representations, the simulation period is extended to the end of August. The simulation flow chart is as follows: the fully-coupled system is run over one month with the two configurations (TerrSysMP(3D) and TerrSysMP(FD)), the two HFDs and five initial conditions of the subsurface. A schematic of the simulation clock is depicted in Fig. 3.5 and will be explained in detail in the following. After the spinup and the construction of the ensemble,
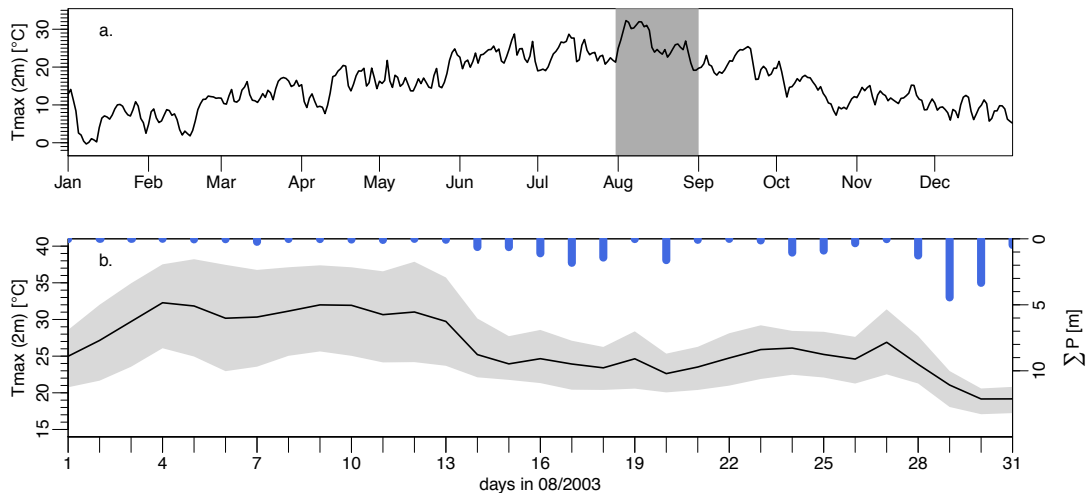
Figure 3.4: (a) Annual time series of the daily maximum 2 m temperature averaged over the PRUDENCE region France from ERA-Interim. The simulation period for the current study is highlighted in gray. (b) Time series of daily maximum $T2M$ [$°C$] and daily aggregated precipitation from ERA-Interim. The shaded polygon indicates the spatial standard deviation over the PRUDENCE region France.

the resulting five initial states are used to initialize the production runs of the fully-coupled system which are performed with two configurations of the subsurface, TerrSysMP(3D) and TerrSysMP(FD) for both HFDs and all initial conditions. In order to keep the simulations of the two configurations and the two HFDs comparable on such a large domain, the atmosphere is re-initialized on a daily basis at 00 UTC (illustrated by the solid blue arrows in Fig. 3.5, impact discussed in Sec. 3.4.1). However, the surface and subsurface states are only initialized once on the 1st of August, and evolve over the entire simulation period without re-initialization. Thus, the experiment setup for both cases differs from a "free forecast run" over one month and moreover delineates a one-day forecast of the atmospheric model COSMO with an underlying transient simulation of the surface and subsurface, assuming that the subsurface is a reasonable estimate of reality.

Furthermore, it is noted that each re-initialization of COSMO at 00 UTC with spatial fields of ERA-Interim leads to an imbalance of the land surface with the atmosphere. The autocorrelation of the 2 m temperature indicates that this imbalance has an effect up to 6 hours (not shown). Thus, the first 6 hours (marked gray in Fig. 3.5) of each day are discarded and only the hours 7-24 of each day (marked blue in Fig. 3.5) are used in the analyses.

3.3.5 Analysis methods

The focus of the analysis is placed on two variables from each compartment: the water table depth (WTD [m]) and the total water storage in the upper 3 m (TWS [m], expressed as equivalent water height) for the subsurface, the sensible (H [W/m^2]) and latent (LE [W/m^2]) heat flux for the land surface, and the 2 m air temperature ($T2M$ [$°C$]) as well as the convective available potential energy ($CAPE$ [J/kg]; formulation can be found in Appendix A) as a vertically integrated stability measure for the atmosphere. Furthermore,
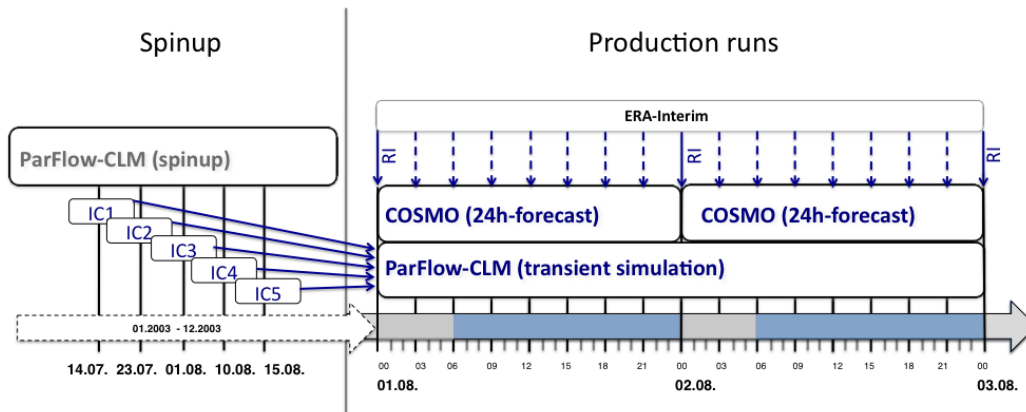
Figure 3.5: Simulation flow chart of the fully-coupled system with daily re-initialization (RI) of the atmospheric model COSMO with ERA-Interim at 00 UTC. Spinup runs illustrate the simulation of Parflow-CLM in order to approach a dynamic equilibrium of the subsurface state. Ensemble simulations of the fully-coupled system are conducted based on 5 different subsurface-land surface states from the spinup simulation.

volumetric soil moisture content (θ $[m^3/m^3]$) and the lifting condensation level (LCL $[m]$) are used to analyze feedback pathways.

The purpose of this study is a sensitivity analysis of the hydrologic and energy cycles to the groundwater representation and the subsurface characteristics. As no calibration has been performed, the simulations do not necessarily closely reproduce observations. Nevertheless, a comparison to ERA-Interim and ERA-Interim/Land is performed in order to stress the different subsurface representations and indicate the comparability to observations and reanalysis products. ERA-Interim/Land is an offline simulation of the improved one-dimensional land surface scheme H-TESSEL from the ECMWF driven with ERA-Interim (Balsamo et al. 2009). Note that, while this product is described as a reanalysis data set, no soil and subsurface parameters are assimilated. Additionally, the offline simulation of ERA-Interim/Land neglects the land surface-atmosphere feedbacks. Nevertheless, ERA-Interim/Land is used to evaluate the land surface states compared to simulations performed with TerrSysMP. The comparison of these systems is performed on the downscaled (DS) ERA-Interim and bilinearly interpolated ERA-Interim/Land fields from 0.5° to 0.11° resolution. We refer to these downscaled fields as ERA-Interim(DS) and ERA-Interim/Land(DS).

The analysis is subdivided into three parts. First, descriptive statistical differences between the simulations with TerrSysMP, ERA-Interim and ERA-Interim/Land are illustrated using density plots. Here, general differences between the two groundwater configurations, the two HFDs, and the ensemble realizations are indicated. Furthermore, average diurnal cycles illustrate the local difference over PRUDENCE regions exemplarily. The responses induced by the two groundwater configurations among all compartments are examined for different classes of WTD using box-whisker-plots. Secondly, the illustrated differences are attributed to feedback pathways using correlations. It addresses both, the water- and energy-cycle related feedbacks and describes the land feedback as the correlation of LE with soil moisture, and the atmospheric segment as the correlation of $T2M$ and LCL (and

equivalently, boundary layer height) to LE and soil moisture. This analysis refers to previous studies and findings from Betts et al. (1996), Seneviratne et al. (2010) and Dirmeyer et al. (2014). The soil moisture-temperature feedback strength is illustrated with scatter plots for each PRUDENCE region. Furthermore, bedrock-to-atmosphere feedbacks are indicated by correlations of subsurface parameters with land surface and atmospheric variables.

Finally, an analysis of variance (ANOVA) is performed in order to extract significant differences between the simulations with TerrSysMP. Similar analyses have been performed in, e.g., Douville (2004) and Bosshard et al. (2013). The methodology of ANOVA is described in detail in Appendix B. The resulting coefficient of determination describes, which proportion of the total variance is explained by a certain treatment. The treatments considered are (i) the groundwater configurations (CONFIG), (ii) the HFDs and (iii) the ensemble realizations (ENSEMBLE), as well as (iv) the temporal variability (TIME). A large value indicates that the model is able to specify the response, given the treatment, or in other words, different treatments lead to different responses. Additionally, the significance of this coefficient is tested with a F-test. Non-significant values indicate that there is statistically no difference between the responses to different treatments (i.e., the different treatments have no effect on the response). Technically, a one-way ANOVA is performed for each treatment, such that the variability of the prescribed treatments in the sensitivity experiment is compared to the large temporal variability. Finally, the ANOVA is performed on the difference between the two groundwater configurations, ΔTerrSysMP(3D-FD), for the different variables (ΔWTD(3D-FD), ΔTWS(3D-FD), ΔLE(3D-FD), ΔH(3D-FD), $\Delta T2M$(3D-FD) and $\Delta CAPE$(3D-FD)) in order to eliminate the strong temporal dynamics over the simulation period and extract underlying processes and patterns. Here, treatments (HFD, ENSEMBLE or TIME), which induce systematic differences between the two groundwater configurations, are determined.

3.4 Results

3.4.1 Impact of daily re-initilization

The experiment design including a daily re-initialization of the atmosphere aims at keeping the general circulations patterns of the performed simulations. Non-local effects, such as different general circulation and advection in the atmosphere could weaken or diffuse the signal of the subsurface effect on atmospheric processes. This effect cannot be completely eliminated, as especially advection may be influenced by land-atmosphere feedbacks. However, a daily re-initialization of the atmosphere with ERA-Interim fields keeps the meso-scale circulation of the performed simulations close to each other and consequently allows for a grid-point comparison.

The simulated differences of pressure and wind speed at selected model levels are comparably small (root mean square differences (RMSD) of $6 - 35\ Pa$, Fig. C.1) and increase towards the planetary boundary layer and the surface, where an increased influence of the subsurface and land surface is expected. Moreover, most of these differences follow subsurface patterns and spatially averaged time series (Figs C.2 and C.3) of wind speed over the
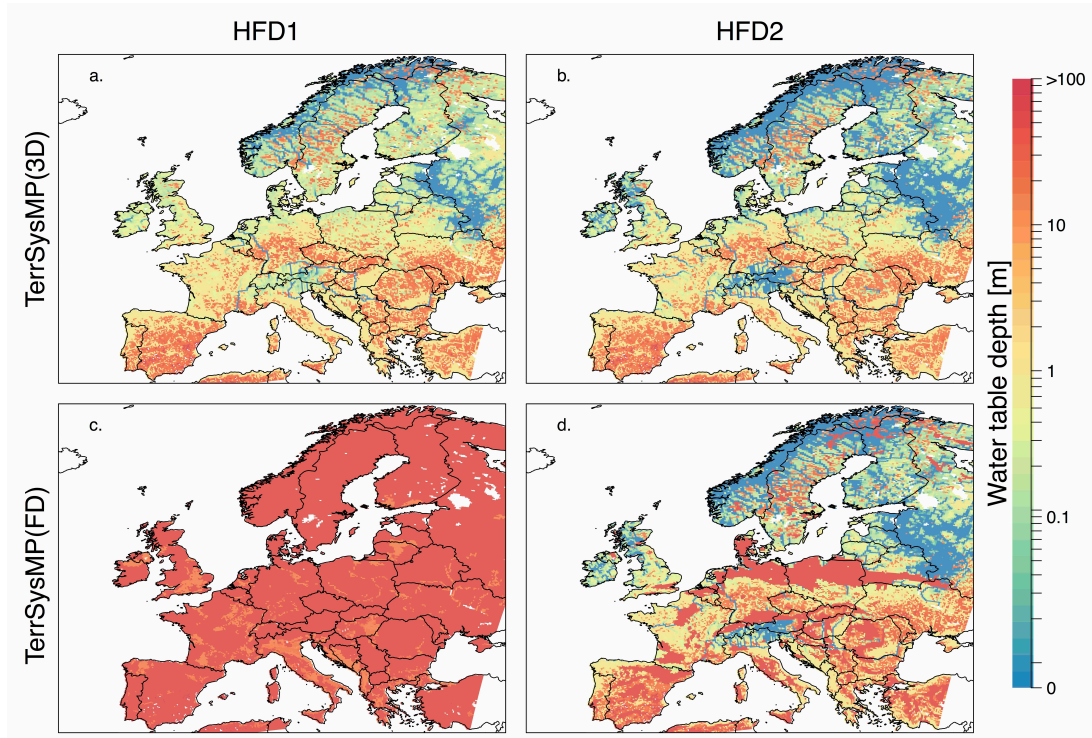
Figure 3.6: Snapshots of the simulated WTDs with TerrSysMP(3D,HFD1), TerrSysMP(3D, HFD2), TerrSysMP(FD,HFD1) and TerrSysMP(FD,HFD2) over the focus domain on 10th of August 2003.

PRUDENCE region illustrate the dependency on subsurface characteristics. Consequently, it is assumed that the general circulation is fairly similar between all simulations. Analyzed differences may thus be mainly caused by local differences arising from the subsurface and land surface states.

3.4.2   Simulated water table depth

Figure 3.6 illustrates a snapshot of TerrSysMP simulated WTD in August 2003. The overall WTDs simulated with TerrSysMP(3D) are comparable to the WTD composition by Fan et al. (2013), showing large-scale patterns, which follow the terrain and a shallow WTD along the coastlines. Furthermore, a shallow WTD is found in arid valleys, thereby suggesting terrain-driven moisture convergence (Fan et al., 2013), but also inundated wetlands in low-land regions, e.g., Netherlands. The simulated WTD emphasizes local gradients, such as the valley-to-ridge-gradient forming rivers, e.g., in the Scandinavian Mountains.

Note that the simulation is performed on a macro-scale, therefore only large-scale over-land flow is simulated. Thus, river corridors are well represented, while flow in river channels is only roughly approximated. The simulated overland flow has a minimum width depending on the horizontal grid resolution, which is usually much larger than any actual river width over the European continent. The respective depth of the overland flow is consequently much smaller than the measured depth of the river flow. The repeated spinup over the extreme dry year 2003 leads to an intermittence of overland flow along these river corridors in Mid-Europe, e.g., along the rivers Seine, Rhone and Loire in France. Still, shallow

*WTD*s along the river channels indicate the physically consistent large-scale convergence of groundwater along river corridors and the memory effect of the soil, e.g., the river channels of the Tajo and the Ebro in Spain. Mountainous regions, such as the Alps are spatially restricted and exhibit a relatively shallow *WTD* because of the low permeability defined for bedrock. This effect is potentially exaggerated through the spatial resolution used, as shown by Sulis et al. (2011) and Shrestha et al. (2015). The simulated *WTD*s with TerrSysMP(3D,HFD1) and TerrSysMP(3D,HFD2) approach each other and are mostly independent of the HFD applied. Consequently, in this setup, the 3D physics are the main driver for the subsurface state. In contrast, simulations with TerrSysMP(FD) are highly sensitive to the HFD used and exhibit very different spatial patterns strongly modulated by the prescribed hydraulic conductivity. Although there exists no physical *WTD* in TerrSysMP(FD), the *WTD* is calculated analogously as the distance from the land surface where the soil is continuously saturated vertically, in order to illustrate the diverging subsurface dynamics. In TerrSysMP(FD,HFD1), the *WTD* drops rapidly and the subsurface dries out quickly, while the *WTD* in TerrSysMP(FD,HFD2) is determined by the hydraulic conductivity in the subsurface. Areas of small conductivities are closer to the physics-based state (as simulated with TerrSysMP(3D)) and areas of high conductivities are closer to the HFD1 scenario and drain out rapidly (Fig. 3.6).

3.4.3 Impact of subsurface hydrodynamics at different space and time scales

Full focus domain and simulation period

On a daily basis, the overall difference between the two configurations, the two HFDs and the individual ensemble realizations is explored using probability density plots. The density functions of individual ensemble realizations and the mean ensemble of daily mean *WTD* and *TWS*, daily maximum *LE* and *H*, daily maximum *T2M* and *CAPE* over all terrestrial grid cells of the focus domain are shown in Fig. 3.7. Overall, the largest differences among all compartments are simulated between TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1) arising from the *WTD* differences between the two HFDs in the FD case, shown in Fig. 3.6. These differences are simulated for all compartments but decrease from the subsurface over the land surface into the atmosphere. The latent (sensible) heat flux from the FD simulations is decreased (increased) compared to the 3D simulations, with minor differences for the HFD2 and large differences for HFD1. Due to the re-initialization, only small differences appear for the atmospheric variables. While the distributions of *CAPE* do not differ significantly (except for the TerrSysMP(FD,HFD1) simulations), a small but clear shift of the distribution of the 2 *m* air temperature towards higher temperatures in TerrSysMP(FD,HFD1) is visible. Overall, the ensemble realizations of each system differ only marginally and almost no difference between the ensemble realizations is visible for the atmospheric variables.

A comparison to the respective distribution of ERA-Interim and ERA-Interim/Land indicates the degree of similarity to the driving atmospheric model and its land surface
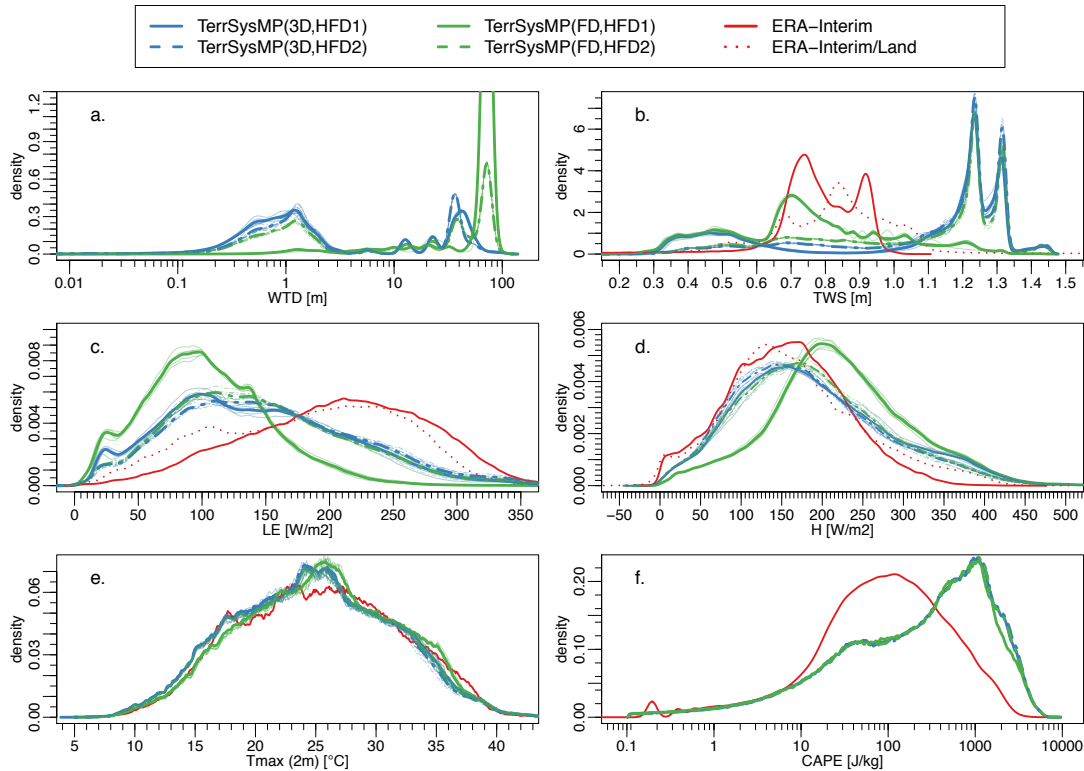
Figure 3.7: Density plots of (a) daily mean WTD [m], (b) daily mean TWS in the upper 3 m [m], (c) daily maximum LE [W/m^2] (d) daily maximum H [W/m^2], (e) daily maximum $T2M$ [$°C$] and (f) daily maximum $CAPE$ [J/kg] over the focus domain for single ensemble realizations of TerrSysMP(3D) (blue lines) and TerrSysMP(FD) (green lines) and ERA-Interim or ERA-Interim/Land (red lines). The thick line shows the respective density of the ensemble mean, and the thin lines show the single ensemble member density. The solid (dashed) lines refer to simulations with HFD1 (HFD2). Note that H-TESSEL in ERA-Interim and ERA-Interim/Land does not simulate dynamic WTDs and that ERA-Interim/Land is driven with atmospheric forcing from ERA-Interim.

model H-TESSEL. For the subsurface and the land surface, large differences between H-TESSEL and TerrSysMP(3D) can be expected mainly due to the subsurface complexity. On the contrary, TerrSysMP(FD) might approach H-TESSEL as both use a 1D subsurface parameterization. Figure 3.7 indicates large differences to ERA-Interim Land, especially for the land surface and the subsurface. While the water storage lies between the water storages simulated with TerrSysMP(3D) and TerrSysMP(FD), the latent heat flux exhibits larger values and the sensible heat flux in turn is slightly smaller but comparable to the distributions of H with TerrSysMP (except for TerrSysMP(FD,HFD1)). However, in both cases, differences might arise from the different resolutions, the different land use and soil definitions (i.e., porosity) in addition to different model parameterizations. Due to the daily re-initialization of the atmosphere, only minor differences of the daily maximum 2 m temperature to ERA-Interim are observed. Figure 3.7 shows that the frequency of temperatures above 35 $°C$ decreases in TerrSysMP simulations, while the frequency increases for temperatures between 23 and 27 $°C$. Nevertheless, the distribution of the daily maximum temperature is quite similar to both versions of TerrSysMP due to the short time after re-initialization. On the other hand, larger values of $CAPE$ are simulated, possibly emerging from different resolutions and convective parameterizations. Note that larger deviations to
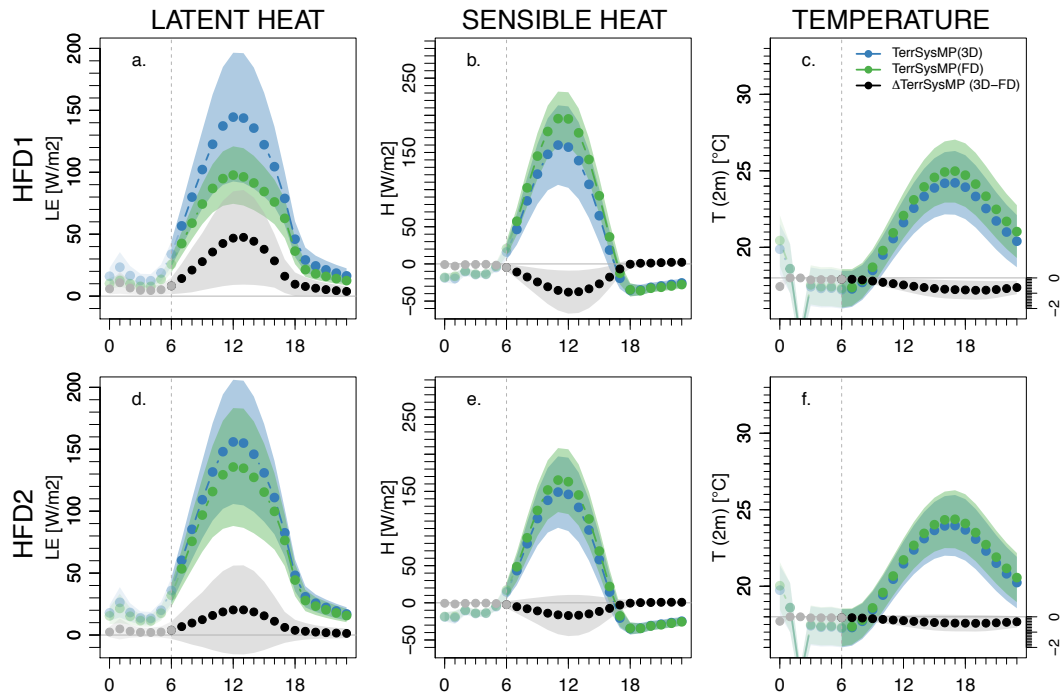
Figure 3.8: Average diurnal cycles of (a,d) LE $[W/m^2]$, (b,e) H $[W/m^2]$ and (c,f) $T2M$ $[°C]$ for TerrSysMP(3D) (blue) and TerrSysMP(FD) (green) over the PRUDENCE region Mid-Europe for the two HFDs ((a,b,c): HFD1 and (d,e,f): HFD2). The black line indicates the difference (3D-FD) and the shaded areas illustrate the standard deviation. Hours 0-6 are disregarded due to the daily re-initialization of the atmosphere and the subsequent imbalance of the fully-coupled system.

ERA-Interim might be expected using transient atmospheric simulations. However, it is not clear yet, whether larger differences between TerrSysMP(3D) and TerrSysMP(FD) can be expected if the atmosphere is not re-initialized on a daily basis. Overall, the density plots indicate large differences for the subsurface, which partly propagate to the land surface, resulting in minor differences for the daily maximum $T2M$ across the full domain and simulation period.

Sub-daily- and daily time scales

In the following, it is shown that the states of the two systems are significantly different on sub-daily- and daily time scales. The average diurnal cycles of LE, H and $T2M$ over the PRUDENCE region Mid-Europe (ME) are illustrated in Fig. 3.8. Significant differences of the average cycle and its variability are simulated for both HFDs between the two groundwater configurations 3D and FD. The largest differences are again simulated in the HFD1 scenario, with up to 50 W/m^2 for ΔLE on average. A smaller difference is visible for H.

Especially the land surface energy fluxes in TerrSysMP(3D,HFD1) exhibit a strong variability, while the variability of TerrSysMP(FD,HFD1) is less pronounced. The differences of both HFDs evolve over daytime and decrease to zero after sunset. Therefore, in the following, the analysis focuses on the daily maximum values of the fluxes and atmospheric
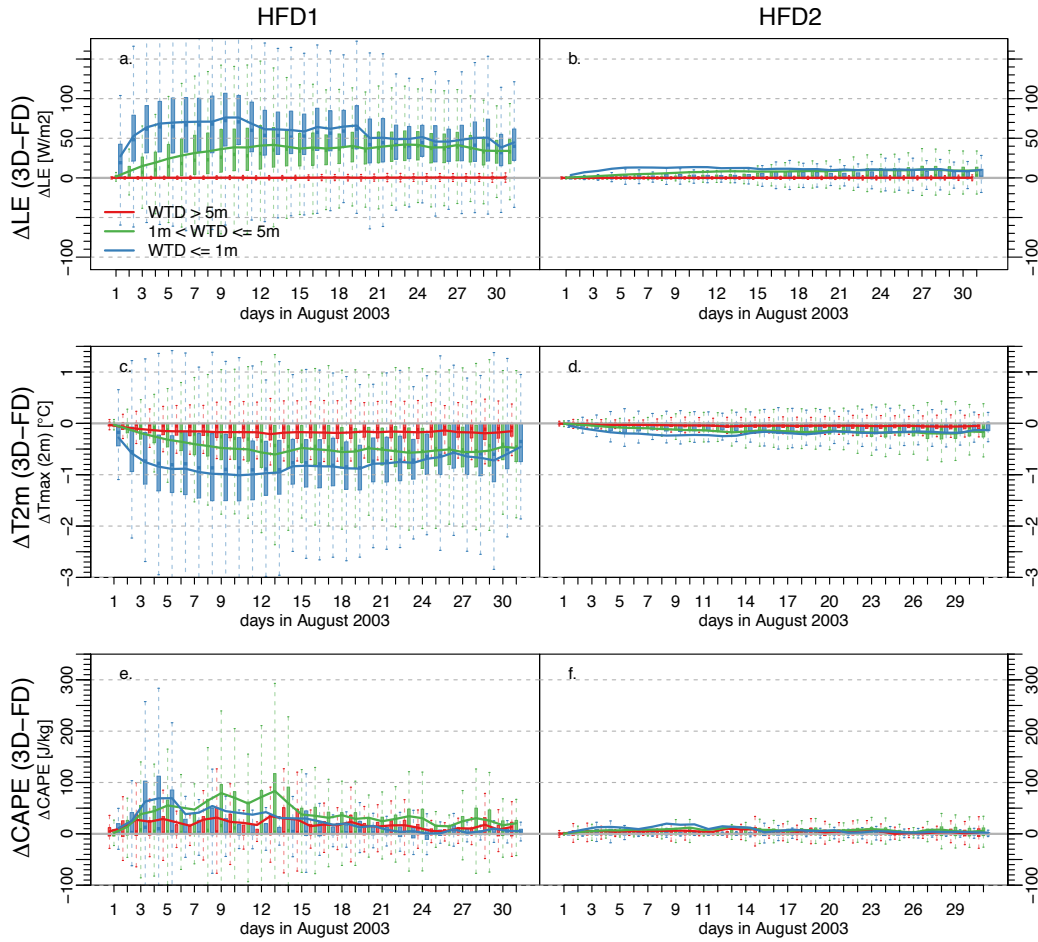
Figure 3.9: Box-whisker plots of daily maximum LE $[W/m^2]$, daily maximum $T2M$ $[°C]$ and daily maximum $CAPE$ $[J/kg]$ differences between TerrSysMP(3D) and TerrSysMP(FD) (i.e. ΔTerrSysMP(3D-FD)) over the entire focus domain conditioned on the water table depth in TerrSysMP(3D): $WTD < 1\ m$ (blue), $1\ m \leq WTD < 5\ m$ (green) and $WTD \geq 5\ m$ (red) for both HFDs ((a,c,e): HFD1 and (b,d,f): HFD2).

variables, which exhibit the largest differences in the simulations. For the subsurface, daily mean values are considered.

The Box-Whisker-plots in Fig. 3.9 show the differences of daily maximum LE, $T2M$ and $CAPE$ over time between the two groundwater configurations 3D and FD for both HFDs over the entire focus domain. The differences are categorized into three classes of WTD according to TerrSysMP(3D): $WTD < 1\ m$ as shallow, $1\ m \leq WTD < 5\ m$ as intermediate and $WTD > 5\ m$ as deep.

Overall, significantly larger differences are simulated for the HFD1 compared to the HFD2 case. Nevertheless, all differences show a clear footprint from the subsurface as the differences strongly depend on the WTD range. The largest differences are simulated for shallow WTDs, and the differences decrease with increasing WTDs. A deep WTD below the root zone has no influence on the land energy fluxes but shows an impact on daily maximum temperatures. The subsurface influence on the tropospheric stability indicated by $CAPE$ is reduced through advection and three-dimensional mixing. Furthermore, the differences in $CAPE$ between the two HFDs corroborate the impact of the groundwater

configuration on the land-atmosphere coupling and the entire atmospheric boundary layer. Overall, large differences in ΔTerrSysMP(3D-FD,HFD1) are obtained through the strong decrease of WTD in TerrSysMP(FD,HFD1); and these differences are less pronounced in ΔTerrSysMP(3D-FD,HFD2) as TerrSysMP(FD) approaches TerrSysMP(3D) in large parts of the domain. This shows that land surface and atmosphere processes, despite the daily re-initialization, are highly sensitive to the HFD in the deeper subsurface. In all cases, the largest differences are observed for shallow WTDs in TerrSysMP(3D), thereby suggesting a strong coupling between all compartments, i.e. the shallow water table, the land surface and the atmosphere.

However, the average difference between the two configurations can also be in the same order of magnitude. Fig. 3.10 is identical to Fig. 3.9 but considers the PRUDENCE region ME only. Here, the differences for HFD1 reach 100 W/m^2 and -1.5 °C on average for shallow WTDs. Compared to the entire focus domain, the average difference in the HFD2 scenario increases and the quartiles show a spatially more variable difference. Compared to the differences of HFD1, this indicates the slower subsurface processes in HFD2. In this case, the free troposphere is less affected as indicated by the smaller $\Delta CAPE$. It is interesting to observe the strong decrease of all differences after the precipitation event at 15th of August. This corroborates the coupling of subsurface hydrodynamics with land surface fluxes and atmospheric dynamics.

3.4.4 Feedback pathways

The differences illustrated in Figs 3.9 and 3.10 arise from feedback mechanisms, which lead to increased temperatures in TerrSysMP(FD), and increased $CAPE$ in TerrSysMP(3D). Specifically, the feedback pathway for the wetter soils in TerrSysMP(3D) is as follows (Pal and Eltahir, 2001): shallow groundwater levels lead to wetter soils, which in turn lead to a decrease in the Bowen Ratio (i.e. an increase of LE) inducing simultaneously a decrease in surface temperatures with an increase of water vapor in the boundary layer. This is commonly associated with a lower boundary layer height and an increase of moist static energy in the boundary layer, leading to a lower LCL and increasing instability (as expressed by $CAPE$). Ultimately, this may lead to moist convection and rainfall. Vice versa, deeper groundwater levels, and thus drier soils, especially in TerrSysMP(FD,HFD1) increase the Bowen Ratio (i.e. decrease of LE, increase of H), induce higher surface temperatures, and ultimately deepen the boundary layer along with an increase of the LCL.

Figure 3.9 and 3.10 indicate that respective differences between TerrSysMP(3D) and TerrSysMP(FD) for LE, H, $T2M$ and $CAPE$, but do not indicate the contributing feedback mechanism with respect to groundwater. In order to attribute these differences to a feedback mechanism, a correlation analysis between groundwater, soil moisture, LE, $T2M$ and LCL is performed. This analysis aims at identifying the dominant feedback pathway and considers bedrock-to-atmosphere feedbacks indirectly, as the impact of groundwater dynamics via shallow soil moisture is examined. Although correlations might not depict the causal connection (Seneviratne et al., 2010), they are indicators for the physics-based coupling relationships and their strength.
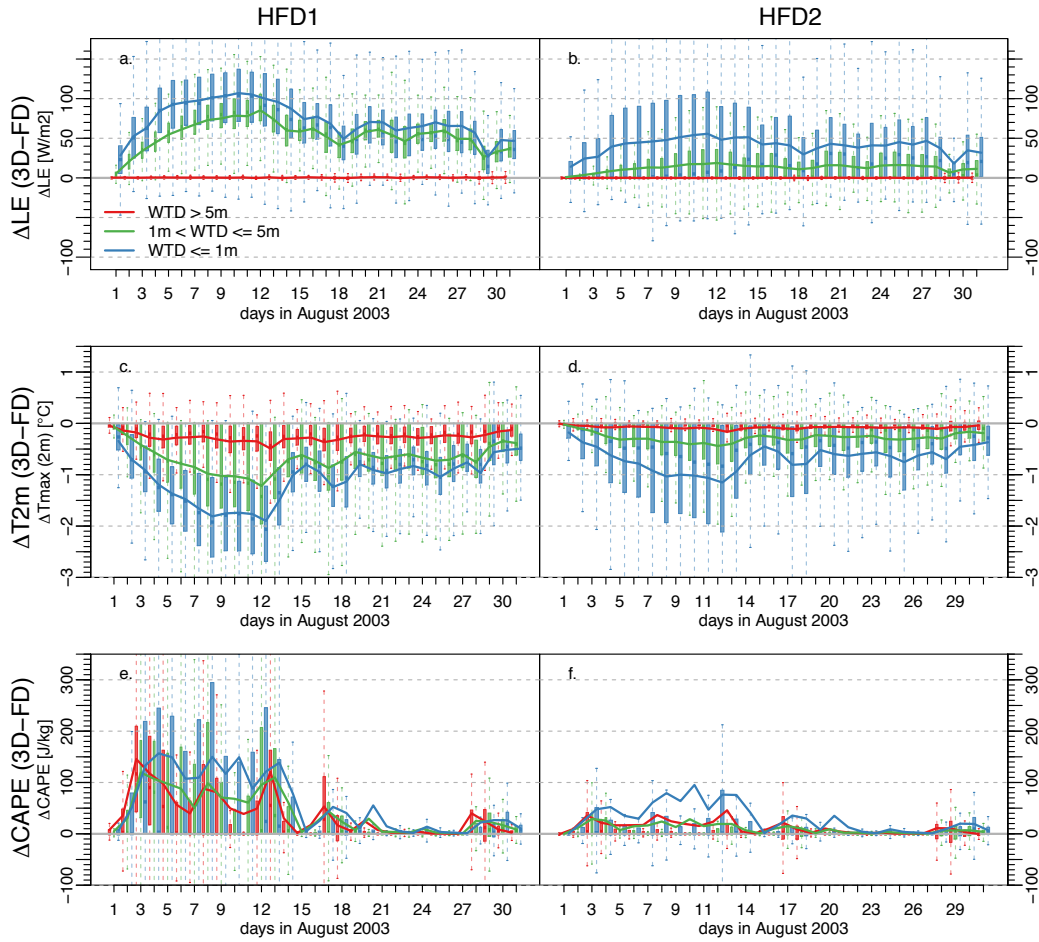
Figure 3.10: As Fig. 3.9, but for the PRUDENCE region Mid-Europe.

Figure 3.11 shows the respective correlation coefficients for the soil-land surface-atmosphere system based on daily values over the first, dry half of the simulation period only, in order to exclude any significant impact of precipitation and emphasize differences between the two groundwater configurations TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1). The correlation between soil moisture and LE (Fig. 3.11a) illustrates the land segment of the water cycle pathway, similar to Dirmeyer et al. (2014), and describes the connection of shallow soil moisture with surface fluxes. The positive correlations in most parts of Europe indicate a positive feedback, except for some de-coupled regions, such as parts of the Iberian Peninsula. Moreover, the correlation between $T2M$ and LE is an indicator for the coupling regime (Seneviratne et al., 2010). A positive correlation indicates the absence of soil moisture-temperature coupling as the atmosphere controls land surface processes through, e.g., limited net radiation (energy-limitation), while strong soil moisture-temperature coupling is characterized by a negative correlation of $T2M$ and LE. In these regions, the land impacts atmospheric processes through soil moisture limitation. Figure 3.11b shows simulated soil moisture-temperature coupling over the heatwave region. Moreover, this coupling is strengthened in TerrSysMP(FD) (Fig. 3.11f) and reveals an intensification of the feedback processes. Figure 3.11g reveals a spatially more uniform feedback, as the correlation between soil moisture and $T2M$ is not masked through the influence of vegetation.
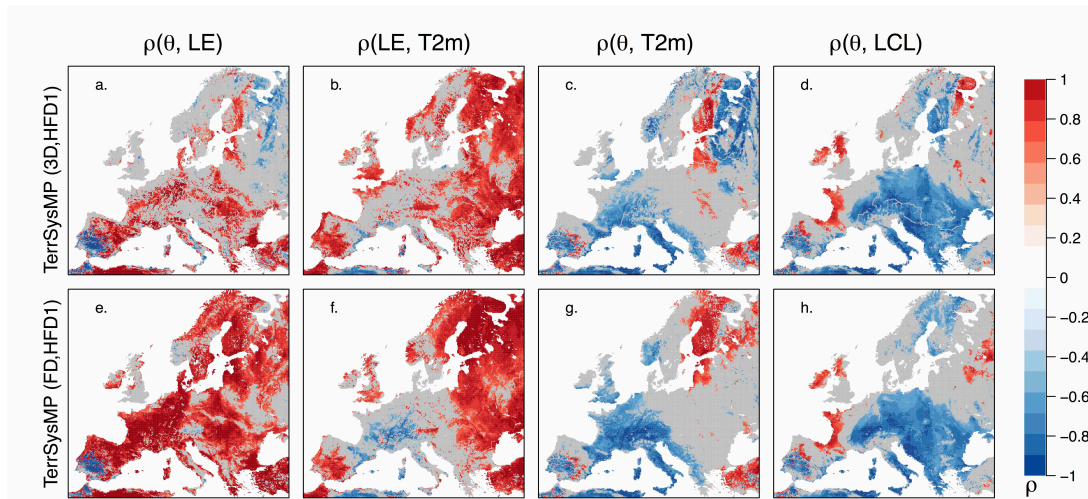
Figure 3.11: Significant Pearson correlation coefficients for TerrSysMP(3D,HFD1) and Terr-SysMP(FD,HFD1) describing the soil moisture-temperature feedback for the first half of the simulation period. Non-significant values are masked grey.

In order to depict not only near-surface processes, the atmospheric feedback pathway is shown as the correlation between soil moisture with the LCL, similar to Betts et al. (1996), Dirmeyer et al. (2006) and Dirmeyer et al. (2014). A negative correlation represents a positive feedback through increasing LCL (and boundary layer height) with decreasing soil moisture and an increasing Bowen Ratio. Large parts of Europe are characterized by this feedback, except for western parts of France and Northern UK. A decoupling (positive correlation) in these regions might be accomplished through the blocking of the high-pressure system during this period and the interference of oceanic moisture. Furthermore, the middle and upper troposphere might be less influenced by land-atmosphere feedbacks due to the daily re-initialization and mixing processes.

Overall, Fig. 3.11 shows a strong soil-land-atmosphere coupling and furthermore a positive soil moisture-temperature feedback for the heatwave region France and Mid-Europe. This does not only affect the near-surface atmosphere, but has an impact on the full boundary layer as indicated by LCL (and previously by $CAPE$). An important finding is that both groundwater configurations show the same feedback mechanisms, but different feedback strengths, which is due to the difference in the representation of groundwater at the continental scale. In TerrSysMP(FD), the coupling is strengthened and results in enhanced feedbacks, which lead to higher temperatures. Thus, the results suggest that the unphysical FD boundary condition for groundwater introduces an overestimation of the coupling between the land surface and the atmosphere, which has been suggested previously by Vautard et al. (2013).

The impact of the soil-land-atmosphere feedback strengths is illustrated in Fig. 3.12 with scatter plots of soil moisture and temperature averaged over the entire simulation period and the respective PRUDENCE regions. Note, each point represents a perturbed initial condition in Fig. 3.12. This analysis is based on the analysis in Fischer et al. (2007)
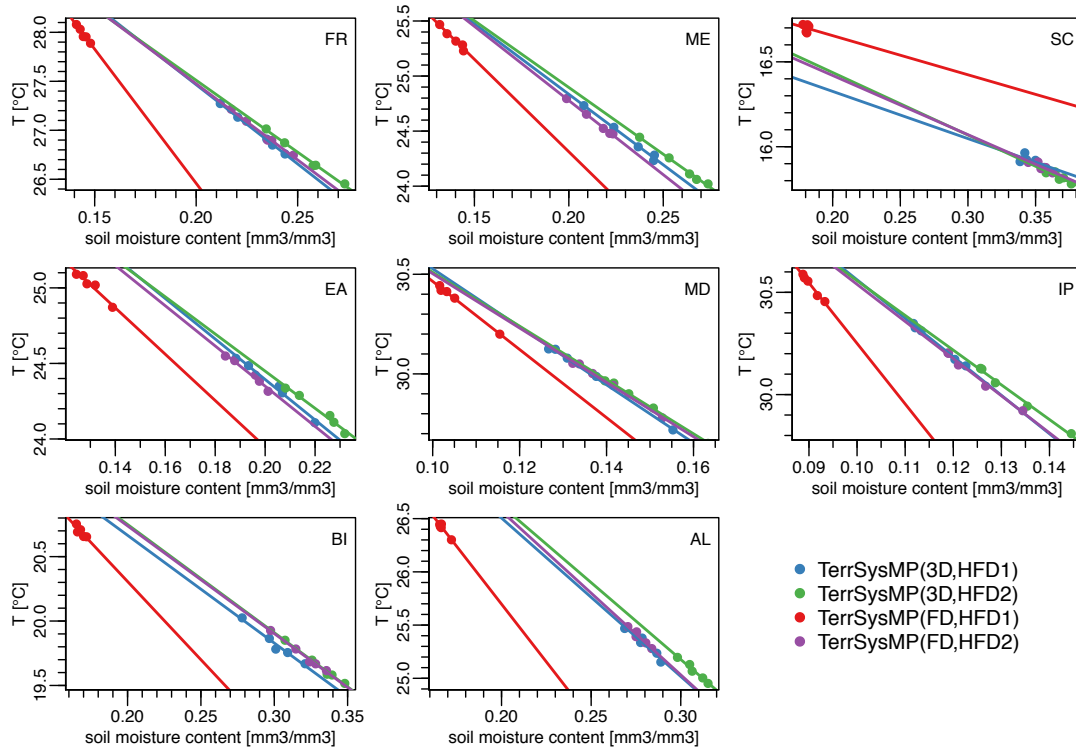
Figure 3.12: Scatter plot of spatially and temporally averaged soil moisture content in the first layer against 2 m maximum temperature for all ensemble realizations and all groundwater representations separated for all PRUDENCE regions (FR – France, ME – Mid-Europe, SC – Scandinavia, EA – Eastern Europe, MD – Mediterranean, IP – Iberian Peninsula, BI – British Islands, AL – Alpine region). The colored lines show the linear regression based on the 5 ensemble realizations arising from different initial conditions.

considering the different setups used in this study. In this plot, the different slopes indicate the different soil moisture-temperature feedback strengths during the simulation period. Interestingly, the realizations from TerrSysMP(FD,HFD1) show the steepest slope for all PRUDENCE regions and hence constitute the strongest feedback with the highest simulated temperatures. The simplified groundwater dynamics in TerrSysMP(FD,HFD1) result in increased temperatures, i.e. TerrSysMP(FD,HFD1) is on average up to +1.5 °C hotter than TerrSysMP(3D,HFD1) over the PRUDENCE region France. However, the differences of the respective slopes for TerrSysMP(3D,HFD1), TerrSysMP(3D,HFD2) and also TerrSysMP(FD,HFD2) are negligible. Thus, groundwater dynamics induce a spatially and temporally more consistent pattern of soil moisture (more moist in river corridors and dryer along the water divides) that is less dependent on subsurface heterogeneity. In case of HFD2, low hydraulic permeabilities result in generally shallow water table conditions and slow drainage over the simulation period, which subdues the effect of the FD boundary condition in TerrSysMP(FD,HFD2).

This suggests that simulations with simplified groundwater physics strengthen the soil moisture-temperature feedback during the heatwave 2003 via shallow soil moisture. This substantially amplifies the feedback mechanism and induces extreme temperatures, significantly higher than in TerrSysMP(3D). However, this feedback strength highly depends on the subsurface characteristics, as prescribed by the HFD.
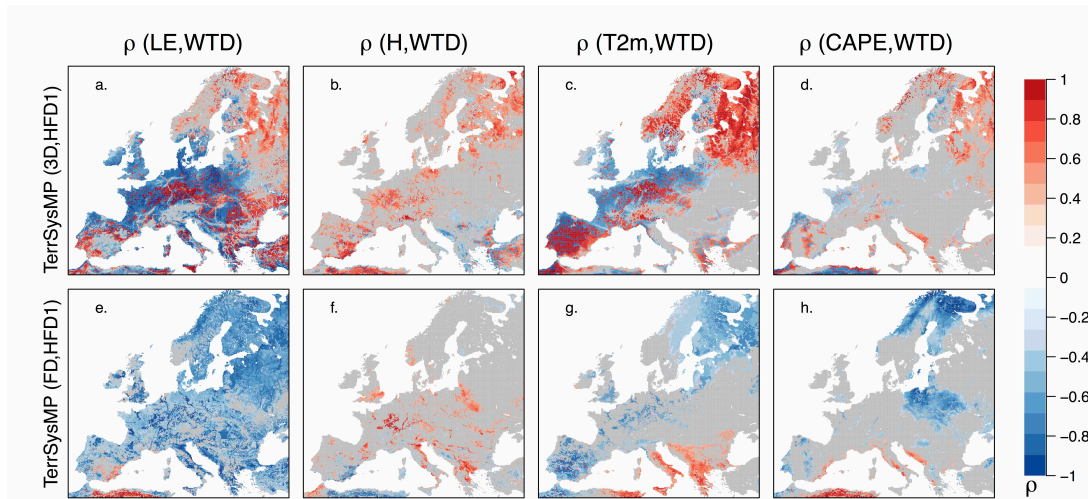
Figure 3.13: Significant Pearson correlation coefficients of daily and ensemble mean WTD with daily maximum LE and H, daily maximum $T2M$ and $CAPE$ for the two groundwater configurations with HFD1 (TerrSysMP(3D,HFD1), TerrSysMP(FD,HFD1)). Non-significant correlations are gray.

In order to attribute the differences to the subsurface-land surface-atmosphere coupling, the correlation coefficients for several subsurface-land surface-atmosphere relations are shown. Figure 3.13 shows the significant correlation coefficient ρ for the daily mean WTD with the daily maximum LE, H, $T2M$ and $CAPE$ for TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1). In general, the correlations between WTD dynamics and heat fluxes, as well as atmospheric variables show a strong connection in the soil-vegetation-atmosphere system. In regions, a positive feedback is depicted through negative ρ of WTD with LE and $CAPE$ and a positive ρ in case of H and $T2M$. This can be explained via capillary rise of water from a shallow water table sustaining LE and thus reducing H in the energy balance. In regions of deeper WTDs in case of TerrSysMP(3D), the positive ρ indicates a decoupling of the deeper subsurface with the land surface (similar pattern in the correlation of WTD with H). Thus, LE is declining due to ambient soil moisture limitation close to the land surface without the effect of capillary rise reducing regional groundwater recharge and thus leading to a constant decline of the water table. Additionally, the physics-based WTD dynamics are positively correlated with the $T2M$, especially in spatially limited heatwave regions (France and Alpine regions).

In case of TerrSysMP(FD), the coupling is increased artificially in all cases. The free drainage boundary condition at the bottom of the soil column leads to strongly increasing WTDs and consequently indicates unphysical correlation patterns. Increased correlation coefficients in Mid-Europe, France and Southern Europe in TerrSysMP(FD) emphasize the strong land-surface-atmosphere feedbacks accelerating the drying of the soil, the increase (decrease) of H (LE) and the increase of $2\,m$ temperatures. The correlation coefficient can even change its sign, e.g., for the correlation between WTD and $T2M$ in Scandinavia. Note that the patterns for HFD2 (not shown) are similar but less pronounced in TerrSysMP(FD) because of slower subsurface dynamics. Overall, the estimated correlations indicate that not only shallow soil moisture, but deeper subsurface processes, such as WTD dynamics,

may influence land surface-atmosphere feedbacks. Yet, these correlations need to be studied over longer time series and further discussed, whether or not these bedrock-to-atmosphere feedbacks have a significant impact and lead to different states across all compartments.
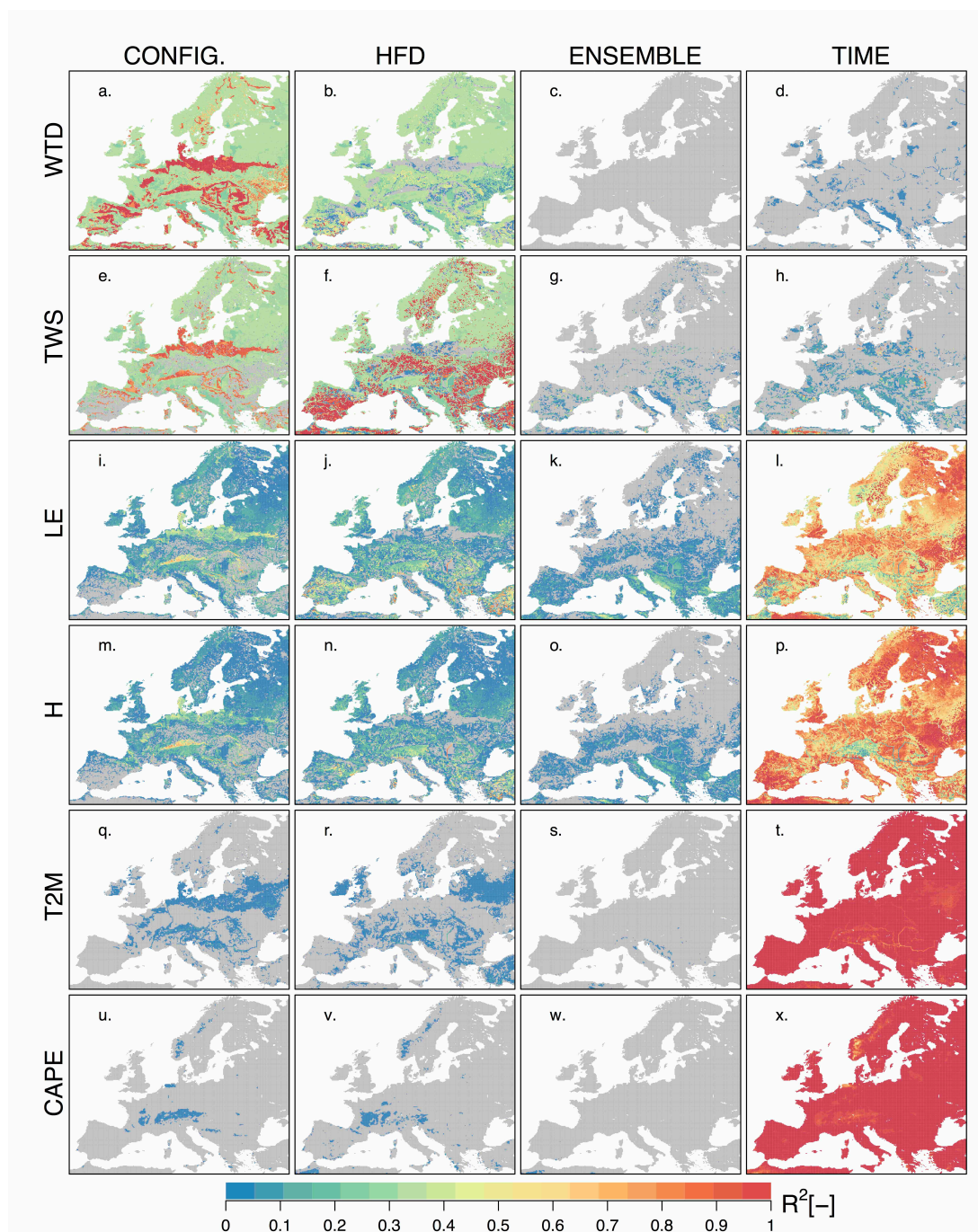

Figure 3.14: Significant coefficient of determination R^2 from a 1-way ANOVA for the different treatments (i) groundwater configurations (CONFIG), (ii) HFDs, (iii) ensemble realizations emerging from different subsurface initial conditions (ENSEMBLE) and (iv) the temporal development in days (TIME), for the variables daily mean WTD and TWS, daily maximum LE, H, $T2M$ and $CAPE$. Non-significant coefficients are gray.

3.4.5 Analysis of variance

In order to determine and extract patterns and variability induced by subsurface hydrodynamics, a one-way ANOVA is performed for several treatments and the respective coefficients of determination are compared. First, the ANOVA is performed on the entire simulation set, including the two groundwater configurations 3D and FD and the two HFDs. The ANOVA is performed for each grid point separately, therefore including the temporal variability of each variable, and exposing spatial patterns within the fully-coupled system. Figure 3.14 illustrates the spatial distribution of the coefficients of determination for the variables (WTD, TWS, LE, H, $T2M$ and $CAPE$) along the columns and for the different treatments (groundwater configurations (CONFIG), HFDs, ensemble realizations (ENSEMBLE) and time (TIME)) along the rows. Large values indicate that a large part of the total variability is induced by the respective treatment.

This analysis shows that the groundwater representation and the HFD in this study have a large influence on the daily variability not only in the subsurface but also at the land surface. Moreover, the atmospheric variability is also partially affected, although the atmosphere is re-initialized on a daily basis. The amount of variability induced by either the groundwater configuration or the HFD decreases from the subsurface to the atmosphere and the atmospheric dynamics are mostly determined by the temporal variability.

In terms of WTD and TWS, the subsurface variability is determined by the two groundwater configurations and the two HFDs implemented. Large values in Fig. 3.14a illustrate diverging WTD dynamics between the 3D and the FD simulations, thereby inducing variability in the land surface and atmospheric compartments. The WTD is mostly determined by the two groundwater configurations, especially in regions of high hydraulic conductivities and for both HFDs. The initial conditions do not influence the total variability of WTD and only some regions exhibit large temporal variability (Fig. 3.14c+d). In contrast, the variability of the TWS in the upper 3 m is influenced by the HFDs in regions with a deep WTD (Fig. 3.14f), while the groundwater configuration (Fig. 3.14e) still determines the variability in regions with a shallow WTD.

The same patterns emerge in the ANOVA for the land surface heat fluxes. While all factors induce variability, the part of variability induced varies spatially. While the variability of LE and H in the Eastern parts of Europe and Scandinavia is determined by the temporal dynamics (Fig. 3.14l+p), the heatwave affected regions, such as France and Mid-Europe are more sensitive to the groundwater configurations (Fig. 3.14i+m), the HFDs (Fig. 3.14j+n) and the initial conditions (ENSEMBLE, Fig. 3.14k+o). Again, regions of shallow WTDs show strong variability induced by the groundwater configurations simultaneously affected by the HFDs. Here, also the initial condition of the subsurface (ENSEMBLE) has a small influence on the total variability of the energy fluxes. Juxtaposition of for the time-induced variability indicates the atmospheric control on H through increasing temporal influence (TIME, Fig. 3.14l+p).

Ultimately, the atmospheric variability in $T2M$ and $CAPE$ is determined by temporal dynamics (TIME, Fig. 3.14t). In other words, the temporal trend during the heatwave in August 2003 is larger than the variability induced by the groundwater representations.

However, the ANOVA indicates a small influence of the groundwater configurations on atmospheric variables, following subsurface patterns. Despite the daily re-initialized atmosphere, the groundwater configurations and the HFDs induce some variability in the atmosphere, especially in regions with a shallow WTD and regions with a strong land-atmosphere coupling (Fig. 3.14q+r). For $CAPE$, the groundwater configurations induce some variability in the southern parts of France and the Alpine region, as being mostly affected during the heatwave 2003 (Fig. 3.14u+v).

Overall, this analysis shows that the groundwater configuration has a significant impact on the subsurface and land surface variability, but the atmospheric variability is mostly determined by the temporal dynamics, such as the long-term trend of the heatwave. Furthermore, subsurface variables and land energy fluxes react highly sensitively to soil hydraulic parameters, as prescribed in the HFDs. Here, the ensemble arising from different initial conditions of the subsurface affects only near-surface and land surface variables, such as heat fluxes.

In order to eliminate the impact of the overlying temporal dynamics and expose underlying patterns induced by the different groundwater configurations, the ANOVA is performed on the difference of the groundwater configurations, ΔTerrSysMP(3D-FD). Again, grid point-differences are analyzed revealing the treatments determining ΔTerrSysMP(3D-FD) on a spatial scale (Fig. 3.15). The structure of Fig. 3.15 is as before, with treatments (HFD, ENSEMBLE and TIME) along the rows and variables (ΔWTD, ΔTWS, ΔLE, ΔH, $\Delta T2M$, $\Delta CAPE$) along the columns. Inspection of Fig. 3.15 in the direction of the different treatments shows that subsurface variability is mostly induced by the HFD and TIME. The spatial patterns are reversed, i.e. HFD dominates ΔTerrSysMP(3D-FD) in regions of low hydraulic conductivity (Fig. 3.3), while TIME dominates ΔTerrSysMP(3D-FD) in high conductivity regions (Fig. 3.15a-c). Interestingly, ΔTWS, ΔLE, ΔH, $\Delta T2M$ and $\Delta CAPE$ exhibit similar patterns for the different treatments suggesting that the groundwater configurations significantly impact land surface and atmospheric variability. This impact is dominated by subsurface heterogeneity (HFD) in low conductivity regions and by internal temporal variability (TIME) in high-conductivity regions across all compartments of the terrestrial system. The initial condition (ENSEMBLE), has no significant impact except on $\Delta T2M$ (Fig. 3.15n). Juxtaposing ΔLE and ΔH for ENSEMBLE (Fig. 3.15h+k) suggests that the grid point response of the heat fluxes is spatially distributed by lateral mixing in the lower atmosphere.
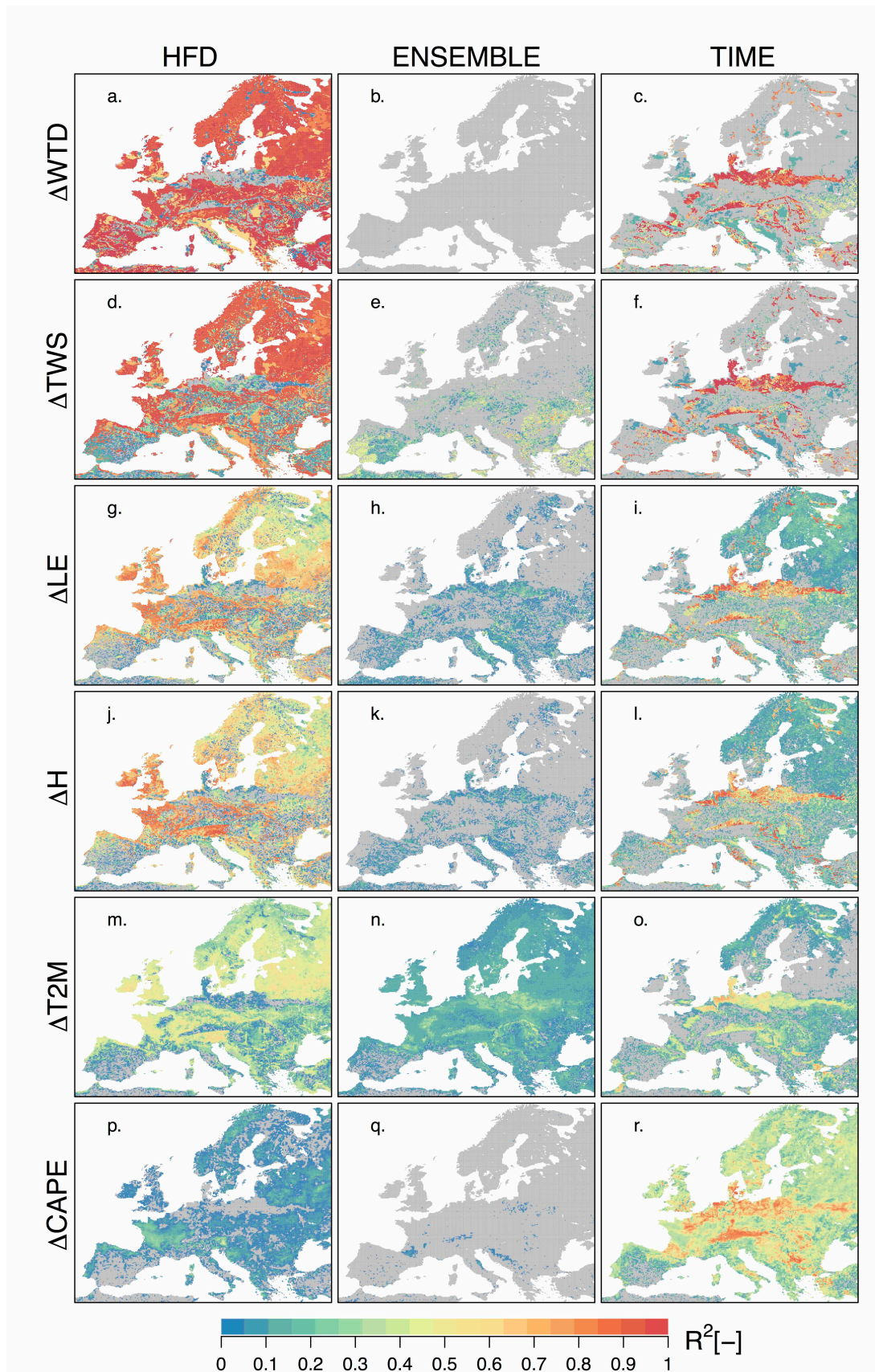
Figure 3.15: Significant coefficient of determination R^2 from a 1-way ANOVA for the difference between the two groundwater configurations (ΔTerrSysMP(3D-FD)) and for the different treatments (i) using different HFDs, (ii) ensemble realizations emerging from different subsurface initial conditions (ENSEMBLE) and (iii) the temporal variability (TIME), for the variables daily mean WTD, TWS, daily maximum LE, H, $T2M$ and $CAPE$. Non-significant coefficients are gray.

3.5   Summary and discussion

The objective of this study is to explore the sensitivity of a fully-coupled soil- vegetation-atmosphere system at the continental scale to groundwater processes. The first hypothesis was that the groundwater representation significantly influences land surface-atmosphere processes during August of the heatwave 2003. This hypothesis was tested with a sensitivity analysis using an integrated fully-coupled soil-vegetation-atmosphere model over the European continent. Here, groundwater dynamics were represented by two different configurations, (1) physics-based three-dimensional groundwater dynamics and (2) a one-dimensional free drainage approach, and two different hydrofacies distributions (HFDs). These representations were set up over the European CORDEX domain and ensemble simulations were conducted, applying different initial surface-subsurface states. The sensitivity analysis showed that, while the physics-based simulations were mostly independent of the HFD, the parameterized one-dimensional free-drainage simulations were highly sensitive to changes in subsurface characteristics. This led to significant differences between the two groundwater configurations but strongly depended on the HFD, as the hydraulic conductivity prescribed in the HFD determined the drainage rate in the free-drainage simulations. The differences between the two groundwater configurations followed subsurface patterns and the largest differences were simulated for shallow WTDs and across all compartments. Not only near-surface variables, such as the $2\ m$ temperature were affected by the groundwater representation but the entire planetary boundary layer and the free troposphere indicated by differences in $CAPE$.

The second hypothesis was, that groundwater dynamics alleviate the evolution of extremes through attenuated land-atmosphere feedbacks during this heatwave. Correlation analysis showed, that a simplified groundwater configuration enhances land-atmosphere coupling via shallow soil moisture, intensifies the feedback mechanism and leads to increased temperatures. Vice versa, groundwater dynamics impede the evolution of such enhanced feedback mechanisms. This analysis is in line with studies from Vautard et al. (2013) and the feedback may be overestimated in the FD case. However, the feedback strength is further dependent on the subsurface characteristics, and only minor differences between 3D and FD were observed for HFD2. In addition, correlation analyses between subsurface dynamics, land surface, and atmospheric dynamics showed strong feedbacks between the compartments for both groundwater configurations including spatially varying and contrary feedback mechanisms. Finally, the statistical difference between the simulations with the fully-coupled system including different groundwater representations was explored using the analysis of variance. In this analysis the variability induced by the groundwater configuration, the HFD or the ensemble arising from different subsurface initial conditions, was compared to the total variability including the time component. It was shown that the groundwater configuration induces variability across all compartments with decreasing impact from the subsurface to the atmosphere. While the ensemble had only a minor impact in dry regions with a deep WTD, the groundwater configurations and the HFD had a significant impact on the variability in the fully-coupled system. In regions with shallow WTDs and a high hydraulic conductivity, the configurations had more influence than the HFD.

Especially for atmospheric dynamics, the variability induced by the groundwater representation was found to be smaller than the temporal variability. An ANOVA on the difference between the two groundwater configurations eliminated this effect and further exposed the propagation of subsurface patterns into atmospheric processes. This analysis showed that the two groundwater configurations lead to significant differences across all compartments but strongly depended on the HFD applied.

A major limitation of this study is the horizontal resolution employed. In general, it is noted that the land surface and subsurface models can (and should) be implemented at higher resolutions. However, this makes the setup of this already extremely complex modeling system (with heterogeneous soil) even more complex, more extensive and more difficult to solve. Thus the current setup of TerrSysMP has, from an atmospheric point of view, the highest resolution that is computationally feasible and allows us to run the fully-coupled system in appropriate wall-clock time in advanced supercomputer environments. While atmospheric processes are well parameterized, especially the hydrologic surface and subsurface processes act at much smaller spatial scales and much longer time scales, yet, the study showed that groundwater dynamics already have an impact at 12 km resolution. Furthermore, the short simulation period of one month and the daily re-initialization limit the generality of the findings, but it is not clear yet, if a less frequent re-initialization leads to increased atmospheric variability induced by subsurface dynamics. Nevertheless, it is a step towards an ESM at the regional scale and furthermore supports high-resolution hydrological modeling by better capturing important feedback processes beyond the simplified parameterizations of surface-subsurface processes in land surface schemes of regional climate models.

Overall, the question arises whether the observed differences remain or even increase over longer time periods in the order of years and decades and in the context of regional climate modeling. Additionally, how do the two systems evolve if the atmosphere is not reinitialized? This study indicates that the two systems develop into two different states and we expect even higher differences, and increased spatial variability further propagating into the atmosphere with the potential of initiating feedbacks for long-term simulations. Yet, these questions remain subject to future research.

Chapter 4

Human water use impacts on the strength of the continental sink for atmospheric water

Abstract

In the hydrologic cycle, continental landmasses constitute a sink for atmospheric moisture as annual terrestrial precipitation commonly exceeds evapotranspiration. Simultaneously, humans intervene in the hydrologic cycle and pump groundwater to sustain, for example, drinking water and food production. Here, we use a coupled groundwater-to-atmosphere modeling platform, set up over the European continent, to study the influence of groundwater pumping and irrigation on the net atmospheric moisture import of the continental landmasses, which defines the strength of the continental sink. Water use scenarios are constructed to account for uncertainties of atmospheric feedbacks during the heatwave year 2003. We find that human water use induces groundwater-to-atmosphere feedbacks, which potentially weaken the continental sink over arid watersheds in southern Europe. This feedback is linked to groundwater storage, which suggests that atmospheric feedbacks to human water use may contribute to drying of watersheds, thereby raising water resources and socio-economic concerns beyond local sustainability considerations.

Key points

- Human water use alters the continental sink for atmospheric water across watersheds.

- Atmospheric feedbacks induced by human water use contribute to drying at the watershed scale over Europe.

- Integrated modeling systems are required to quantify anthropogenic impacts on the water cycle through groundwater-to-atmosphere feedbacks.

4.1 Introduction

Groundwater sustains ecosystems and simultaneously secures global water and food supply via, for example, groundwater abstraction and irrigation (Taylor et al., 2013; Wada et al., 2012). Groundwater pumping significantly contributes to declining water tables in aquifers and watersheds around the world (Famiglietti, 2014; Rodell et al., 2009; Scanlon et al., 2012; Long et al., 2013). Previous studies also indicated that the effect of human water use (HWU) on the terrestrial hydrology is comparable to climate change effects at the regional scale (Ferguson and Maxwell, 2012) and may already exceed sustainability at the global scale (Jaramillo and Destouni, 2015).

Irrigation may decrease land surface temperatures through increasing evapotranspiration (ET) and associated surface cooling (Kueppers et al., 2008; Lobell et al., 2009) and mitigate heat extremes (Thiery et al., 2017), thereby potentially counteracting warming effects due to climate change (Hirsch et al., 2017). Simultaneously, groundwater depletion may limit future irrigation water and hence increase the heatwave risk (Lu and Kueppers, 2015). Via ET, HWU may also impact precipitation (P) at different spatial scales and contribute to sea level rise (Wada et al., 2016b). Previous studies linked observed P changes over the continental US to the development of groundwater pumping and an intensification of irrigation during the last century (Moore and Rojstaczer, 2001; Alter et al., 2015; DeAngelis et al., 2010). Simulations indicated that irrigation can either increase P but also inhibit P initiation locally (Barnston and Schickedanz, 1984; Douglas et al., 2009; Pei et al., 2016), and beyond the local impact, may also affect terrestrial P in remote regions through changes of the atmospheric moisture transport (de Vrese et al., 2016). These P feedbacks can in turn impact river flow at remote locations (Wang-Erlandsson et al., 2017) and affect water resource governance (Keys et al., 2017). Consequently, the impacts of HWU on water resources may impact sustainability especially during dry and hot summers (such as the European heatwave in 2003) in which water demands increase significantly (e.g., van der Velde, 2010) and land-atmosphere feedbacks are expected to be strongest (Fischer et al., 2007; Ferranti and Viterbo, 2006, Miralles et al., 2014). While such feedbacks cannot be easily unraveled using observations (e.g., Turner et al. 2007; Karl and Trenberth, 2003), recent progress in Earth system modeling allows to perform sensitivity studies in order to identify potential feedback pathways.

However, previous atmospheric impact studies applying regional climate models assume an unlimited supply of water at the land surface to mimic irrigation or use simplified hydrologic models, which do not close the terrestrial hydrologic cycle from groundwater into the atmosphere (Wada et al., 2017; Nazemi and Wheater, 2015). Thus, the connection of groundwater with the land surface energy balance and atmospheric processes has not been taken into account explicitly (Anyah et al., 2008; Gilbert et al., 2017), and their impact on subsurface water storages is not considered. In turn, hydrologic studies including groundwater processes are just beginning to emerge at the continental scale (Bierkens et al., 2015; Maxwell et al., 2014; Maxwell and Condon, 2016), and still neglect atmospheric feedbacks to HWU.

In this study, we perform a sensitivity study by applying a mass and energy conservative groundwater-to-atmosphere modeling platform in an ensemble approach, unifying the atmospheric and hydrologic paradigms. We study how HWU, here considered as groundwater abstraction and irrigation, may alter the strength of the continental moisture sink. The analysis covers different spatial and temporal scales across European watersheds during the European heatwave in 2003. While this heatwave was an extreme event in the recent past, droughts and heatwaves are expected to increase in the future (e.g., Beniston, 2004; Perkins et al., 2012; Beniston and Stephenson, 2004), being tightly coupled to the projected water demand, as indicated by, e.g., Schewe et al. (2014) and Wada et al. (2013). Yet the full impacts of HWU during extended dry periods remain poorly understood constituting a challenge for water security and water resource vulnerability (Vörösmarthy et al., 2000; Oki and Kanae, 2006).

4.2   Methods

The integrated Terrestrial Systems Modeling Platform (TerrSysMP, Shrestha et al., 2014; Gasper et al., 2014) is used to perform this sensitivity study. TerrSysMP consists of the atmospheric model COSMO (Baldauf et al., 2011; Doms and Schättler, 2002), the Community Land Model (Oleson et al., 2008) and the surface-subsurface model ParFlow (Kollet and Maxwell, 2006; Jones and Woodward, 2001), coupled through OASIS-MCT (Valcke, 2013). TerrSysMP accounts for fully-coupled atmospheric and land surface-subsurface hydrologic processes, and incorporates a realistic representation of groundwater dynamics, the hydraulic connection to surface water, and two-way feedbacks with atmospheric processes. The platform has already been set up over the European CORDEX domain at 0.11° resolution (Keune et al., 2016) and only small modifications to the setup are made (Appendix D). In this study, we incorporate HWU as groundwater abstraction and irrigation in the surface-subsurface flow model of TerrSysMP (Appendix D). In order to account for uncertainty of the land-atmosphere feedbacks and P initiation, four HWU scenarios are constructed and results are based on the comparison of the HWU ensemble to a natural reference model run (NAT) covering the heatwave year 2003 over the European continent. Here the natural reference run is representative of an undisturbed water cycle without groundwater pumping and irrigation. The HWU scenarios are based on realistic estimates of daily groundwater abstraction for industrial, domestic and agricultural use, as well as irrigation (HWU1, Wada et al., 2012; Wada et al., 2016a, and HWU2, Siebert et al., 2010; Siebert and Döll, 2010; Figs D.2- D.4). In addition, two water use schedules are applied for each data set (HWU1-1 and HWU2-1: daytime water use; HWU1-2 and HWU2-2: nighttime water use). In order to keep the large-scale atmospheric circulation as close as possible to the driving ERA-Interim reanalysis, we apply the spectral nudging approach (von Storch et al., 2000) for the horizontal wind components above the planetary boundary layer. This technique arrives at P events closer to reality also reducing biases (Appendix D). Simulations start from a multi-year spinup of the hydrological compartments of TerrSysMP.

This set of simulations is used to examine groundwater-to-atmosphere feedbacks of HWU on the strength of the continental moisture sink, measured by the atmospheric divergence

$div(Q)$. We define the continental sink of moisture as $C_{SI} = -div(Q) = P - ET$, if P exceeds ET, $P > ET$. Vice versa, the continent is a net source of water to the atmosphere, $C_{SO} = div(Q) = ET - P$, if ET exceeds P, $ET > P$. Subsequently, we identify the impact of this change of the continental sink on continental drying as simulated by changes in the subsurface storage S and runoff R (Fig. D.1). A detailed description of the land-atmosphere water balance and the analyzed differences induced by HWU is provided in Appendix D.

4.3 Results and discussion

4.3.1 Water use impacts on the continental sink

At the annual time scale and in the natural hydrologic cycle, most of the continental land masses are a net sink of water as P exceeds ET (C_{SI} in Fig. 4.1a). Hence the atmosphere is a net importer of water for the continent, which is well-known (Gimeno et al., 2012). However, there exists spatial and temporal variability depending on atmospheric dynamics, land cover and groundwater dynamics. Regions and entire watersheds can also be a net source of moisture to the atmosphere seasonally and for warm summer months (C_{SO} in Fig. 4.1b-c), but also for the full year in areas with strong groundwater convergence, such as the Pannonian Plain and the Transylvanian Plateau (Fig. 4.1a). In those cases, the atmosphere exports more water of continental origin to other regions, increasing the potential for P recycling (Gimeno et al., 2012; Brubaker et al., 1993; Keys et al., 2016).
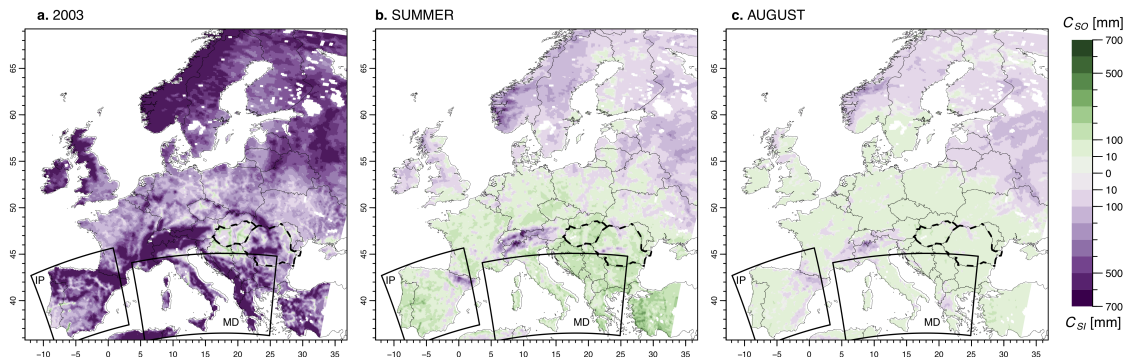
Figure 4.1: Strength of the natural continental sink-source. Continental sink-source relationship for atmospheric moisture simulated with NAT over (a) the full year 2003, (b) Summer: June-July-August (JJA) 2003, and (c) the heatwave month August 2003. Purple shaded areas indicate that the land is a net sink for atmospheric moisture (C_{SI}: $P > ET$, $div(Q) < 0$), green shaded areas indicate that the land is a net source of moisture to the atmosphere (C_{SO}: $ET > P$, $div(Q) > 0$). The boxes indicate the PRUDENCE regions Iberian Peninsula (IP) and Mediterranean (MD). Hungary and Romania, in which the Pannonian Plain and the Transylvanian Plateau are located, are highlighted by thicker dashed lines.

HWU perturbs the natural land and atmospheric water balances and may amplify or attenuate the continental sink of water in a consistent manner through systematic differences in $div(Q)$ by changes of evapotranspiration (ΔET) and precipitation (ΔP) (Appendix D). All simulations in this study indicate a decrease of the net sink of moisture, C_{SI}, with respect to the natural reference simulation for the entire continent caused by HWU (between 0.7 $mm/year$ and 3.2 $mm/year$; Table D.1). This decrease in C_{SI} on the order of

$\sim 1 \; mm/year$ is small and the relatively large range of differences indicates high uncertainty. However, the sign of the signal is consistent and its magnitude is comparable to previous studies, which focused on the influence of groundwater pumping on sea level rise (Konikow and Kendy, 2005; Wada et al., 2016b). Inspecting the terrestrial feedbacks of moisture transport, we observe a systematic spatial and temporal redistribution of water due to HWU. Our results indicate that HWU consistently alters the annual continental sink C_{SI} of atmospheric moisture over the most arid watersheds in southern Europe (Fig. 4.2). Regionally, irrigation and pumping and resulting groundwater-to-atmosphere feedbacks lead to a decrease of C_{SI} ranging from 2.6 to 7.8 $mm/year$ (depending on the water use scenario and relative to a natural sink of 456 $mm/year$) over the Iberian Peninsula, of which the Guadalquivir basin shows the largest decrease ranging from 13.2 to 19.2 $mm/year$ (with $C_{SI(NAT)} = 251 \; mm/year$). France also shows a decrease in C_{SI} for 3 out of the 4 water use scenarios ranging from 1.7 to 12.3 $mm/year$ (with $C_{SI(NAT)} = 374 \; mm/year$), where the Seine basin shows the largest consistent decrease in C_{SI} as a result of decreased P (Fig. D.7). A net increase in C_{SI} is simulated over the Mediterranean (by 0.7 to 12.3 $mm/year$, with $C_{SI(NAT)} = 493 \; mm/year$), but a large number of small watersheds also experience a P deficit (Fig. D.7) accompanied by an increase in ET (Fig. D.8). This indicates the spatial contrast in the Mediterranean climate regimes, which are influenced by the Apennines mountains in the western regions of the Italian Peninsula and the Mediterranean suboceanic climate in the east.

The simulated feedbacks and sensitivities to HWU and soil moisture perturbations are strongest in summer (Wei et al., 2016), which is the main season for irrigation and pumping (Fig. D.4). In summer, HWU can increase the continental source, C_{SO}, by more than 25 mm for the managed basins in the southern parts of the Iberian Peninsula (Fig. 4.2e-l). While the southern regions of the Iberian Peninsula are characterized by an increasing export of atmospheric moisture, mainly through increased ET as a result of irrigation (Figs D.2- D.3 and Fig. D.8), the northern regions (i.e., Ebro and Douro) act as both, attenuated sinks or attenuated sources driven by the uncertainty of the P response in August (Fig. 4.2i-l and Fig. D.7). At the daily time scale, HWU may even turn an entire watershed from a natural sink to a human induced source of moisture to the atmosphere (Fig. D.6). Thus, the continental sink of the two most intensively water managed regions in Europe (Figs D.2-D.3) appears to be altered systematically by HWU and the consequences strongly depend on the atmospheric feedback processes. Many arid watersheds in southern Europe import less atmospheric moisture influencing continental freshwater storage through complex feedback pathways, thereby raising water resources and sustainability concerns. These sustainability concerns are addressed in a section below.
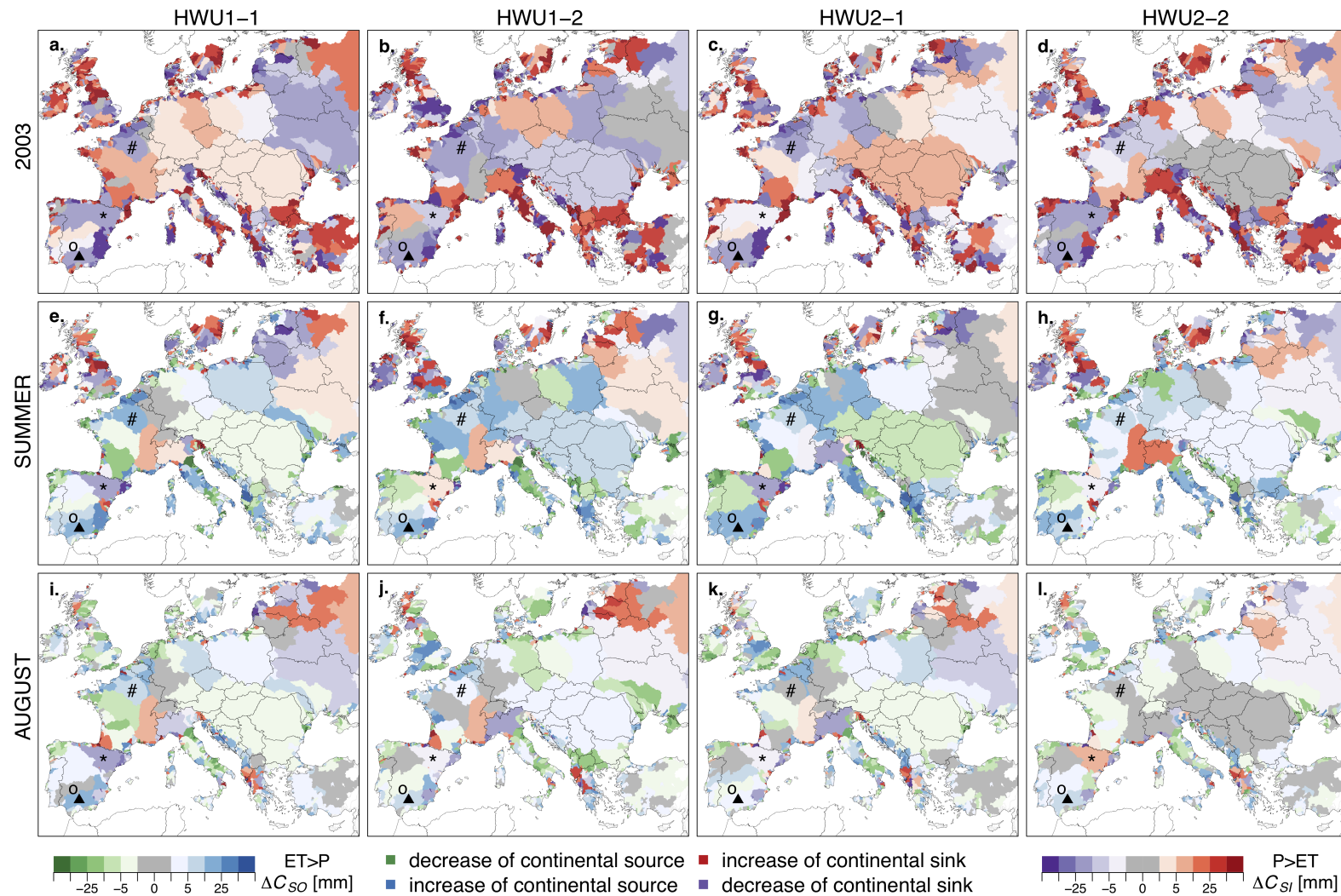
Figure 4.2: HWU-induced alteration of the continental moisture sink. Mean HWU-induced increase or decrease of the continental source or sink of moisture for atmospheric water, compared to the natural reference run (i.e. HWU-NAT) over European watersheds for all water use scenarios along the columns (HWU1-1 (a,e,i), HWU1-2 (b,f,j), HWU2-1 (c,g,k), HWU2-2 (d,h,l)) over the full year 2003 (a-d), summer 2003 (JJA, e-h), and August 2003 (i-l). The watersheds Ebro ($\star$), Guadalquivir (Δ), Guadiana (o) and Seine (#) are marked.

4.3.2   Consistency across space and time scales

While the ensemble simulations in this study show large variability and, hence, large uncertainty of the atmospheric feedbacks to relatively small perturbations of soil moisture by HWU, with respect to magnitude and position, all water use scenarios show considerable consistency of the simulated feedbacks across various space- and time scales.

Figure 4.3 shows $\Delta(div(Q))$ from HWU1 and HWU2 along the x- and y-axes, respectively. Increased percentages of regions, watersheds, and grid points falling in the upper right and lower left quadrant of Fig. 4.3 indicate consistency between the water use scenario simulations and a systematic feedback signal. Deviations from the 1:1 line illustrate the large variability of the simulated feedbacks, which are expected due to the natural chaotic dynamics of the terrestrial system and the high non-linearity of the feedbacks. Local feedbacks at the grid point scale exhibit large spatial variability for all water use scenarios with similar spatial patterns (Fig. D.5), but indicate a consistent feedback signal (approximately 69% for the year and 74% for summer). This variability may be caused by differences in the irrigation and groundwater abstraction rates and in the location between the different water use data sets, as well as less-pronounced atmospheric feedbacks during nighttime water use. The results show that uncertainty decreases and consistency increases with increasing watershed size (Fig. 4.3b). Moreover, small-scale perturbations can also trigger consistent changes in the hydrologic cycle at the regional scale (Fig. 4.3a and b). Our simulations indicate a consistent alteration of the hydrologic cycle for the Iberian Peninsula and the Mediterranean for all water use scenarios, with an increase of the continental sink over the Mediterranean and a decrease of the continental sink over the Iberian Peninsula. Simulated feedbacks for other regions are uncertain, including central Europe, France, eastern Europe and the Alps (Fig. 4.3a), which were less irrigated and experienced less groundwater pumping in 2003 compared to southern Europe (Figs D.2- D.3, Wada et al., 2014).

Moreover, inspection of Fig. 4.3 along the columns reveals that consistency increases between the water use scenarios towards summer and single summer months, and is strongest in the summer season and the heatwave month of August (Fig. 4.3g-l). Interestingly, the P feedbacks often exceed the ET feedbacks due to non-linear interacting processes, yet indicate a stronger consistency for August 2003 (Figs D.7 and D.8). Here, all water use scenarios show that HWU triggers convection and lead to a simulated increase in P in southern France and parts of mid-Europe (Fig. D.7), thereby potentially alleviating water stress and consequently the heatwave, as shown in other irrigation-based studies (Thiery et al., 2017).

4.3.3   Impact on continental water resources

Systematic changes in the continental sink-source relationship for atmospheric water due to HWU are contributing to changes of continental freshwater storage and discharge through complex feedback pathways; that is, HWU can lead to a decrease of continental freshwater storage through an increase of ET, a decrease of P, an increase of discharge, and their combination, respectively (a detailed description of the land-atmosphere water balance is
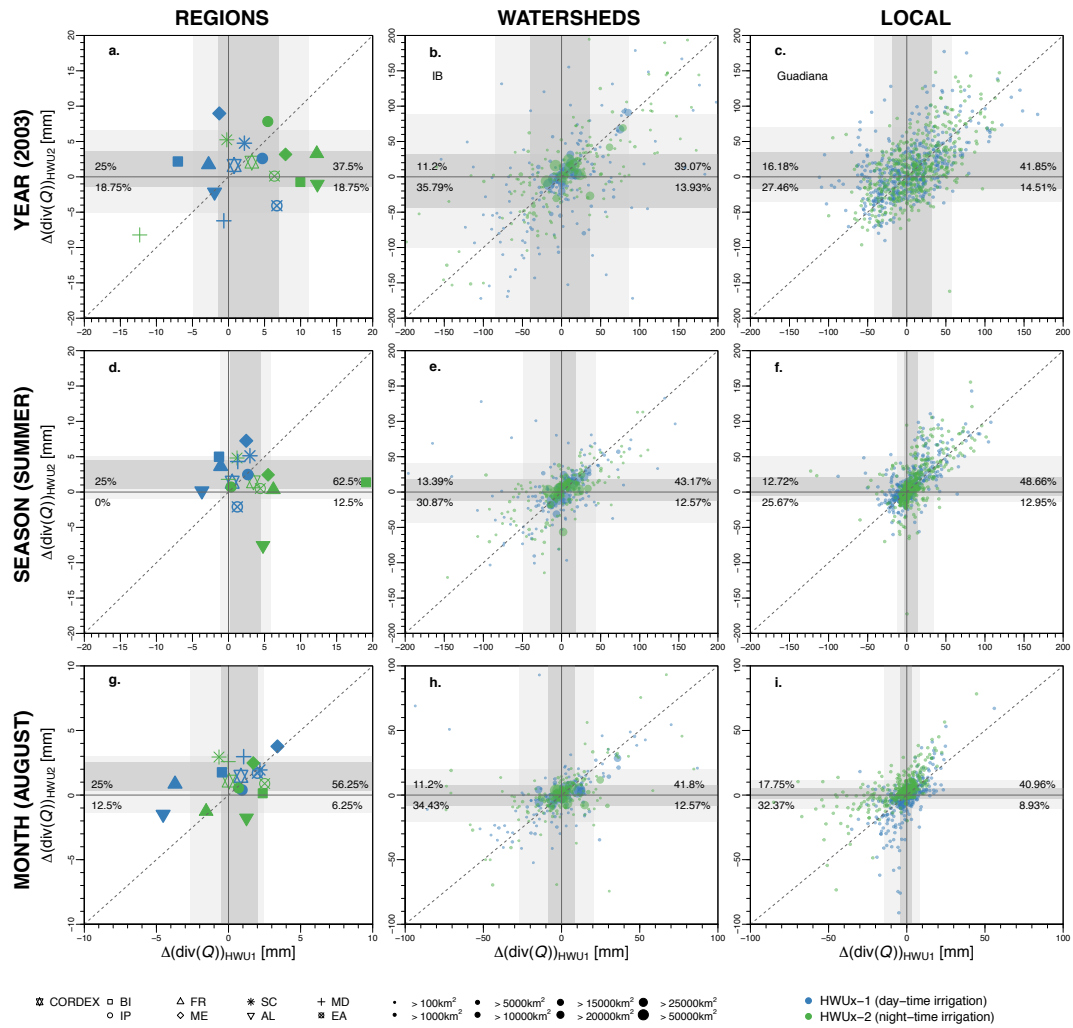
Figure 4.3: Consistency between the water use scenarios (HWU1 and HWU2) induced feedbacks of atmospheric divergence across spatial and temporal scales. Consistency of $\Delta(div(Q))$ over a range of spatial and temporal scales: from year to season to month (along the rows), and from PRUDENCE regions to watersheds to grid points (along the columns). Figures 4.3(b,e,h) illustrate exemplarily the consistency of watersheds over the Iberian Peninsula, and (c,f,l) grid point scale feedbacks in the Guadiana basin. Grey shaded areas show the 10, 25, 75 and 90%-quantiles, respectively. Blue (green) points indicate simulations using daytime (nighttime) water use. In (b,e,h), the size of the symbols is commensurate with the watershed size. The symbols indicate the PRUDENCE regions British Islands (BI), France (FR), Scandinavia (SC), Mediterranean (MD), Iberian Peninsula (IP), Mid-Europe (ME), Alps (AL) and Eastern Europe (EA). Note that the scales of each subplot are different. The consistency between the water use scenarios is indicated by the percentages in the lower left and upper right quadrant.

provided in the Appendix D). These feedbacks can be identified and quantified from our simulation results constituting a major advantage of the integrated groundwater-to-atmosphere modeling approach.

We find a clear inverse relationship between the atmospheric feedbacks on the change of the continental moisture sink ΔC_{SI} and the subsurface water storage changes ΔS, as illustrated in Fig. 4.4; that is, a decrease of the continental sink for atmospheric moisture leads directly to a decrease in continental freshwater storage. Note that ΔS was calculated over the full subsurface column from the bottom of the aquifer to the land surface including
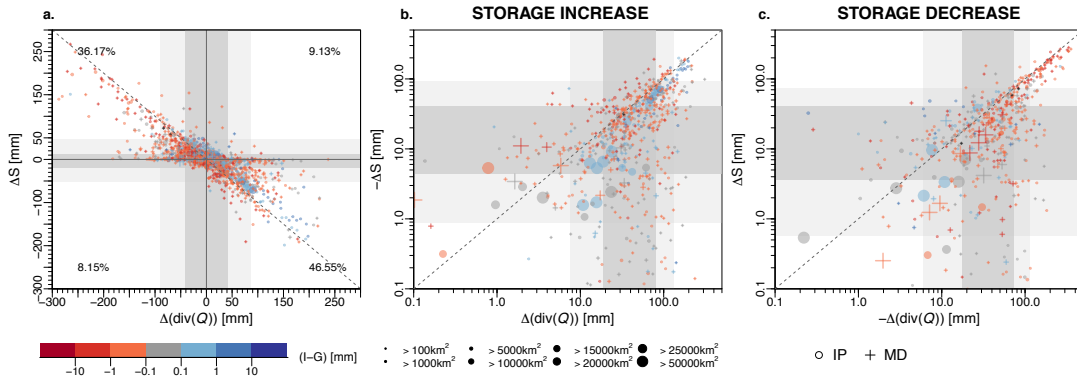
Figure 4.4: Relationship between HWU induced atmospheric feedbacks and soil drying. Annual subsurface water storage changes as a function of changes in the continental sink C_{SI} for all watersheds and all water use scenarios over the Iberian Peninsula and the Mediterranean. $\Delta S < 0$ indicates a subsurface storage decline with respect to the natural simulation NAT over the full year 2003. The grey shaded areas indicate the 10, 25, 75 and 90%-quantiles, respectively. Colors indicate the net HWU as the difference between irrigation (I) and groundwater abstraction (G). (a) displays the entire range, while b (c) shows the relation where an increase (decrease) of the continental sink leads to an increase (decrease) of the subsurface water storage. All units are mm per unit watershed area. The size of the symbols is commensurate with the watershed size.

the variably saturated zone. Thus, while ΔS over some watersheds may be small, localized groundwater storage changes may still be relevant due to transient effects. Approximately 46% of all watersheds plotting in the lower left quadrant of Fig. 4.4a experience a decrease in the strength of the continental sink, while about 36% of all watersheds plotting in the upper right quadrant experience an increase. Figures 4.4b and c show the upper right and lower left quadrant, respectively, in double logarithmic scale for better inspection. Watersheds close to the 1:1 line are of particular interest and raise socio-economic concerns related to water resource sustainability, as increased atmospheric moisture demand induced by HWU is entirely met by a decrease in terrestrial water storages. Moreover, our results indicate that the large watersheds with a deep water table over the Iberian Peninsula and the Mediterranean region are most likely to suffer a water table decline due to HWU and the integrated atmospheric feedbacks (Fig. 4.4c and Figs D.9 and D.10), thereby emphasizing sustainability concerns. Watersheds plotting above and below the 1:1 line respond to changes in the continental sink by changes in groundwater divergence $\Delta div(Q_g)$ (including cross-watershed flow) and continental discharge ΔR (Appendix D). Figure 4.4 illustrates that in most cases, the terrestrial water storage change ΔS is determined by the atmospheric feedbacks ΔC_{SI} to HWU, rather than the net effect of groundwater pumping and irrigation.

4.4   Summary and discussion

The simulations in this study show that HWU alters the atmospheric moisture transport and the strength of the continental sink-source for atmospheric water, and hence leads to nonlocal effects beyond the watershed scale. Our analyses indicate consistent impacts for

watersheds in southern Europe, in which HWU induced feedbacks decrease the strength of the continental sink for atmospheric moisture, C_{SI}, constituting the main contribution to terrestrial water storage declines for a large number of watersheds. The Iberian Peninsula and the Mediterranean encompass the most intensively water managed regions in Europe and show the strongest, consistent alteration of C_{SI} in our simulations during the European heatwave in 2003. However, uncertainty remains for large parts of Europe indicating that even small rates of local HWU in a relatively water-rich region may have an impact on moisture transport and terrestrial hydrology, ultimately contributing to the uncertainty of climate change signals (Allen and Ingram, 2002). These findings suggest that the effect of a decreasing continental moisture sink potentially amplifies water scarcity in regions of strong groundwater abstraction and irrigation, possibly leading to increasing drought and heatwave frequency, strength, and duration. Thus, the simulation results do not corroborate that HWU counteracts an intensification of the hydrologic cycle through climate change effects. This study confirms that HWU is a significant source of uncertainty in global and regional climate simulations not taken into account in current climate model inter-comparison and also impact studies (Jaramillo and Destouni, 2015). Unfortunately, observations of the terrestrial system are prohibitively scarce and error prone especially with regard to the subsurface hydrology to perform this type of study at the continental scale based on measurements. However, in future, available observations must be merged with models via data assimilation approaches to obtain a best estimate of the current state of the water cycle under HWU conditions. We presented the sensitivity of a fully-coupled bedrock-to-atmosphere system for an extreme and dry year. While this might be an upper estimate of the atmospheric feedbacks to HWU under current conditions, the effect might aggravate under global change with increasing occurrence of droughts and heatwaves (Beniston and Stephenson, 2004).

Chapter 5

Potential added value of incorporating human water use on the simulation of evapotranspiration and precipitation in a continental-scale bedrock-to-atmosphere modeling system: A validation study considering observational uncertainty

Abstract

Human activities, such as human water use, have been shown to directly influence terrestrial water fluxes and states. Simulations of soil moisture, river discharge, evapotranspiration and groundwater storage are significantly improved, if human interactions, such as irrigation and groundwater abstraction are incorporated. Yet, improvements through the incorporation of human water use on the simulation of local and remote precipitation are rarely studied, but may contribute to the skill of land surface fluxes. In this study, we evaluate the impact of human water use on the skill of evapotranspiration and precipitation in a fully-coupled bedrock-to-atmosphere modeling platform. The results show that human water use can potentially increase the skill of the simulations across scales. However, observational uncertainty at the watershed scale limits the identification of model deficiencies and added value related to human water use. Locally, daily precipitation statistics potentially benefit from the incorporation of human water use. Although the incorporation of human water use does not remove the wet bias, it can increase the model skill.

> **Key points**
>
> - Gridded observational data sets of evapotranspiration and precipitation exhibit a large uncertainty at the watershed scale.
>
> - Model biases at the watershed scale are subject to observational uncertainty.
>
> - Incorporation of human water use improves the simulation of evapotranspiration and precipitation at the local scale.

5.1 Introduction

The ability to understand, monitor and predict the terrestrial water cycle under current and global change remains a major scientific and socioeconomic challenge (Wagener et al., 2010; Wood et al., 2011). Climate change scenarios along with an expected intensification of the water cycle with more frequent droughts and floods (Huntington, 2006) add uncertainty to future scenarios of water storages and flows (Vörösmarty et al., 2000). Commonly, the intensification is often measured at the interface between the land surface and atmosphere through combined changes in evapotranspiration (ET) or precipitation (P) as an indicator for drying, which has been discussed controversially (Greve et al., 2014). While the assessment of human impacts on the terrestrial water cycle remains challenging (Oki and Kanae, 2006), studies demonstrated that human water management may lead to an intensification of the terrestrial water cycle, especially during droughts (e.g., He et al., 2017; Wanders and Wada, 2015).

Several attempts across multiple disciplines have been made to improve the simulation of the terrestrial water cycle. From an atmospheric perspective, these range from increasing resolutions to improve the simulation of P, e.g., from general circulation models (GCMs) to regional circulation models (RCMs) (Di Luca et al., 2015; Rummukainen, 2016) towards convection permitting simulations (Hohenegger and Schär, 2007; Prein et al., 2013; Prein et al., 2015) to replace simplified parameterization schemes. Yet, especially RCMs with parameterized convection have been identified to exhibit a wet bias (Rojas et al., 2011; Casanueva et al., 2016), which additionally limits their ability to simulate extremes and heatwaves (Weisheimer et al., 2011; Vautard et al., 2013). In this case, a scale separation in space is beneficial to identify the added value, which is often located at the meso-β scale ($20 - 200\ km$; Feser, 2006).

An improved representation of land surface states can further improve the realism of the simulated terrestrial water cycle. The incorporation of uncertainties in land surface models also leads to an improved predictability of P (Orth et al., 2016) and extremes, such as the heatwave over Europe in 2003 (MacLeod et al., 2016). This in turn might indicate the lack of information and realism, that current land surface models comprise (Davin et al., 2016; Tuinenburg and de Vries, 2017; Miralles et al., 2018). Even observation-based estimates of terrestrial ET exhibit relatively large uncertainties (Mueller et al., 2011; Mueller et al., 2013; Trambauer et al., 2014; Sörensson and Ruscica, 2018). Therefore, terrestrial systems models, and data sets to inform these models have been continuously expanded, incorporating additional processes of the water, energy and bio-geochemical cycles including human activities under the assumption that an increasing number and realism of simulated processes of the terrestrial water cycle leads to improved predictive skill (Sulis et al., 2018). One such process is the incorporation of human management of land cover, land use, and water resources. The most commonly studied management practice is irrigation, which may significantly influence P (DeAngelis et al., 2010; de Vrese et al., 2016), and its incorporation can improve the skill of GCM simulations (Thiery et al., 2017).

Large uncertainty in the simulation of the terrestrial water cycle also arises from the subsurface state. From the hydrologic perspective, the simulation of the terrestrial water

cycle can be improved by increasing resolution (Bierkens et al., 2015) and the incorporation of human water use (HWU) beyond irrigation (Wada et al., 2017). Hydrologic model validation comprises a wide range of studies, which evaluate the performance of the simulated terrestrial water states and fluxes, such as soil moisture, ET and discharge, from watersheds to continents (e.g., Döll et al., 2003; Haddeland et al., 2006b; Rakovec et al., 2016; Liu et al., 2017). While the incorporation of water management including groundwater abstraction and reservoirs, has been shown to improve the skill of simulating hydrological extremes (Veldkamp et al., 2018), a high uncertainty of water fluxes simulated with land surface and hydrological models comes from the driving atmospheric forcing and P (e.g., Biemans et al., 2009; Müller-Schmied et al., 2014; Müller-Schmied et al., 2016; Wang et al., 2016).

In an effort to move towards a process-based, and potentially more realistic model based representation of the terrestrial water cycle, studies attempt to bridge the gap between hydrologic and atmospheric research through the incorporation of a better representation of subsurface and land surface hydrology including a HWU interface in Earth System Models (ESMs; Nazemi and Wheater, 2015; McDermid et al., 2017; Pokhrel et al., 2015; Voisin et al., 2013; Hazenberg et al., 2016). While progress has been made in understanding the terrestrial water cycle, the validation of such integrated modeling systems and the identification of the origin of biases becomes increasingly challenging (Oreskes et al., 1994) due to the lack of continuous and co-located observations across the Earth system from groundwater across the land surface into the atmosphere and observational uncertainty (Kotlarski et al., 2017). However, especially for water resource assessments, the identification of biases is essential. In a coupled modeling system, this bias might not only propagate from the local to the regional scale (Addor et al., 2016), but also from one compartment to the other (groundwater to land surface, land surface to atmosphere, and vice versa; Sulis et al., 2017), which potentially aggravates the bias through feedback processes.

There is a growing number of observations including satellites, but also sensor networks and single stations to measure states and fluxes of the terrestrial water cycle (McCabe et al., 2017). Yet, validation studies commonly neglect the fact that observations include additional uncertainty. There exist multiple sources that introduce uncertainty to observations. For example, observational devices may exhibit malfunctions, miscalibration, and location biases. Analogously, all observations have a footprint, i.e., a spatiotemporal reference, for which they are assumed to be representative. The footprint of point measurements is typically very small and on the order of a few meters, especially with respect to small scale processes, such as ET, which is mainly influence by local features, and convective P. On the contrary, satellite images have a much larger footprint on the order of ten to hundreds of kilometers. To bridge the gap between these scales, point measurements are often aggregated to gridded observational products, which introduce an additional uncertainty through assumptions of the underlying algorithm and are especially critical for regions with a sparse coverage of *in-situ* stations.

In this study, we evaluate the integration of HWU in simulations using the continental-scale integrated Terrestrial Systems Modeling Platform (TerrSysMP), a fully-coupled soil-vegetation-atmosphere modeling system, during the heatwave 2003 in Europe. The main objectives of this study are to 1) evaluate the accuracy of integrated groundwater-atmosphere

simulations at the land-atmosphere interface considering observational uncertainty, and 2) assess potential added values of incorporating HWU on simulation skills with a focus on ET and P. The first objective aims at a general evaluation of the agreement and the accuracy of the simulations at the watershed scale at monthly to annual time scales using gridded observational data sets of ET and P. Multiple observational data sets of ET and P are used for validation. Both variables are evaluated separately, but also combined to assess the accuracy and uncertainty of the net water flux at the land-atmosphere interface. The second objective focuses on the added value of HWU in the simulation of terrestrial water fluxes addressing the questions whether and how the consideration of HWU improves the skill of integrated bedrock-to-atmosphere simulations. Additional daily *in-situ* measurements of P and ET from co-located stations are utilized to discuss potential benefits of incorporating HWU.

The outline of this manuscript is as follows. Section 5.2 provides the methodology, the simulations and the experimental design, and introduces the observational data sets and metrics used to evaluate the accuracy and the skill of the simulations. Section 5.3 illustrates and discusses the results for a natural reference simulation. The added value of incorporating HWU is assessed in section 5.4. Summary and conclusions are provided in section 5.6.

5.2 Data and methods

5.2.1 Simulations

In this study, the simulations were carried out with the Terrestrial Systems Modeling Platform (TerrSysMP; Shrestha et al. 2014; Gasper et al., 2014), which was setup over the European CORDEX domain at 0.11° resolution (Keune et al., 2016). In a later version, HWU was considered as groundwater abstraction and irrigation (Keune et al., 2018). A detailed description of the setup can be found in aforementioned references, but a brief description is repeated here.

Setup of the modeling system

TerrSysMP simulates the full terrestrial hydrologic cycle from groundwater to the atmosphere including feedbacks between shallow groundwater, land surface ET and boundary layer processes along with P initiation. Lateral/3D surface-subsurface flow is simulated using ParFlow (Jones and Woodward, 2001; Kollet and Maxwell, 2006; Maxwell, 2013), which is coupled to the land surface model CLM3.5 (Oleson et al., 2008) and the atmospheric weather prediction model COSMO (version 5.1, Baldauf et al., 2011; Doms and Schättler, 2002) through the Ocean Atmosphere Sea Ice Soil coupling tool OASIS3-MCT (Valcke, 2013). In this setup, COSMO was run with a 60 s timestep, while ParFlow and CLM were run with a 180 s timestep. Coupling between the models was performed every 180 s, constituting a very high coupling frequence to account of non-linearities.

COSMO was initialized and driven at the boundaries with the reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF), ERA-Interim (Dee et al., 2011).

In order to keep the large scale atmospheric circulation of all simulations consistent with the reanalysis, spectral nudging (von Storch et al., 2000) was applied. Vertical winds (u,v) were nudged above the planetary boundary layer (p < 850h Pa) with a nudging coefficient of $\alpha = 0.05$ for wavenumbers smaller than 14. The δ-two-stream approximation of the radiative transfer equation according to Ritter and Geleyn (1992) was used. Convection was parameterized with the Tiedtke mass flux scheme (Tiedtke, 1989), and vertical turbulent diffusion was simulated with a 2.5-level closure scheme. Cloud water, cloud ice, rain and snow were simulated with a bulk-water continuity model.

The land surface characteristics for CLM3.5 were based on previous studies (Keune et al., 2016) including an updated land cover and LAI data set based on the Moderate Resolution Imaging Spectroradiometer (MODIS, Friedl et al., 2002). Subgrid-scale heterogeneity was accounted for by a maximum of 4 plant functional types per grid cell.

ParFlow was set up with the hydrofacies distribution 2 (HFD2) from Keune et al. (2016) in order to represent vertically heterogeneous soil and hydrogeologic characteristics. This HFD consists of a soil section (reaching a depth of 3 m) and a deeper subsurface and bedrock hydrogeology (reaching a depth of 103 m). The soil texture was prescribed using the Food and Agriculture Organization (FAO) soil database (FAO, 1988) for the upper ten soil layers. Hydrogeology was prescribed at the lower 5 soil layers using the Gleeson database (Gleeson et al., 2011a, 2011b). To mimic the ocean, a lateral Dirichlet boundary condition with a hydrostatic profile of a shallow water table was applied.

Water use scenarios

In this study, a set of 5 simulations, carried out over the heatwave year 2003, was analyzed. These include a natural reference simulation (NAT) and 4 HWU scenarios. The HWU scenarios were constructed as follows: Two HWU data sets and two water use schedules were applied. HWU1 describes the HWU, i.e., irrigation and (total) groundwater abstraction, from Wada et al. (2012) and Wada et al. (2016a). HWU2 is based on HWU from Siebert et al. (2010) and Siebert and Döll (2010). In both cases, groundwater abstraction accounted for irrigation, and domestic and industrial demand, where domestic and industrial demand from Wada et al. (2012, 2016a) were added to groundwater abstraction estimates for irrigation from Siebert et al. (2010) and Siebert and Döll (2010) for consistency. Irrigation and pumping were applied simultaneously between 06:00 and 18:00 UTC (daytime, HWU1-1 and HWU2-1) or between 18:00 and 06:00 UTC (nighttime, HWU1-2 and HWU2-2). Both data sets were bilinearly interpolated to the 0.11° grid and rates of each daily estimate were added to the top soil layer (irrigation) or subtracted from the bottom soil layer (groundwater abstraction/pumping) at every time step in ParFlow using source/sink terms. All simulations were started from a multi-year spinup over the year 2003 after which a dynamic equilibrium was reached.

In this study, the simulations evaluated include the four water use scenarios (HWU's) and the natural reference simulation (NAT). Note that originally the water use ensemble was not created to represent real-world conditions, but to induce a relatively large spread in order to assess the uncertainty of land-atmosphere feedback related to HWU. Nevertheless, these

simulations are used to assess the accuracy of each water use scenario but also the overall accuracy of the water use ensemble mean (mean of all four HWU's) and test improvements against the natural reference simulation. The water use ensemble mean is calculated using the equally weighted water use scenarios, i.e.,

$$HWU - ENS = \sum_{i=1}^{M} w_i \cdot HWU_i \qquad (5.1)$$

with w_i=1/4 for all scenarios and M=4. Thus, 6 simulations overall (4 water use scenarios, 1 water use mean, 1 natural reference run) are evaluated using observations.

Simulation period

Simulations were performed for the heatwave year 2003, which was characterized by extreme dry and hot conditions over Europe (Schär and Jendritzky, 2004), due to numerous hydro-meteorological factors (Black et al., 2004). An anomalously anticyclonic high over Europe from May to August lead to reduced clouds and P, which in turn lead to desiccating soils. These effects induced a pre-heatwave at the beginning of July, and a mega heatwave in August. Land surface conditions played a crucial role during both heatwave periods and, due to the absence of sufficient soil moisture to meet evaporative demands, were characterized by a heat accumulation through sensible heat. The mega heatwave period was influenced by a steady anticyclonic high over France, which promoted the dominance of the local heat balance, which, in turn, lead to an exacerbating soil moisture-temperature feedback loop and extreme temperature records (Miralles et al., 2014). The exceptional combination of all feedbacks makes the heatwave of 2003 difficult to simulate with current models (Weisheimer et al., 2011). In particular, Weisheimer et al. (2011) showed that it requires not only an improved land surface hydrology, but also improved radiation and convection parameterizations to improve model skill in simulating the heatwave conditions. Here, we assess how the integration of HWU improves the model skill.

5.2.2 Observations

Two types of observations are used for validation. First, gridded observations are used to evaluate monthly, seasonal and annual estimates of ET, P and the net moisture flux $ET - P$, at the watershed scale. All gridded direct observation-based data sets of ET and P that are available for 2003 at a spatial resolution of at least 0.5° and a monthly time scale, are used for validation in order to assess observational uncertainty. These include only diagnostic data sets, which are either interpolated *in-situ* observations, merged observational data sets, and in case of ET, are based on the Penman-Monteith equation or empirically derived formulations, such as the Priestley-Taylor equation. We refrain from using modeled estimates, in which ET and P are estimated with the help of a land surface model and/or a reanalysis, because these do not consider HWU explicitly (e.g., irrigation; Tuinenburg and de Vries, 2017). The gridded observations comprise five P and three ET data sets, which have been used in previous studies (Müller et al., 2011; Greve et al., 2014). We also combine
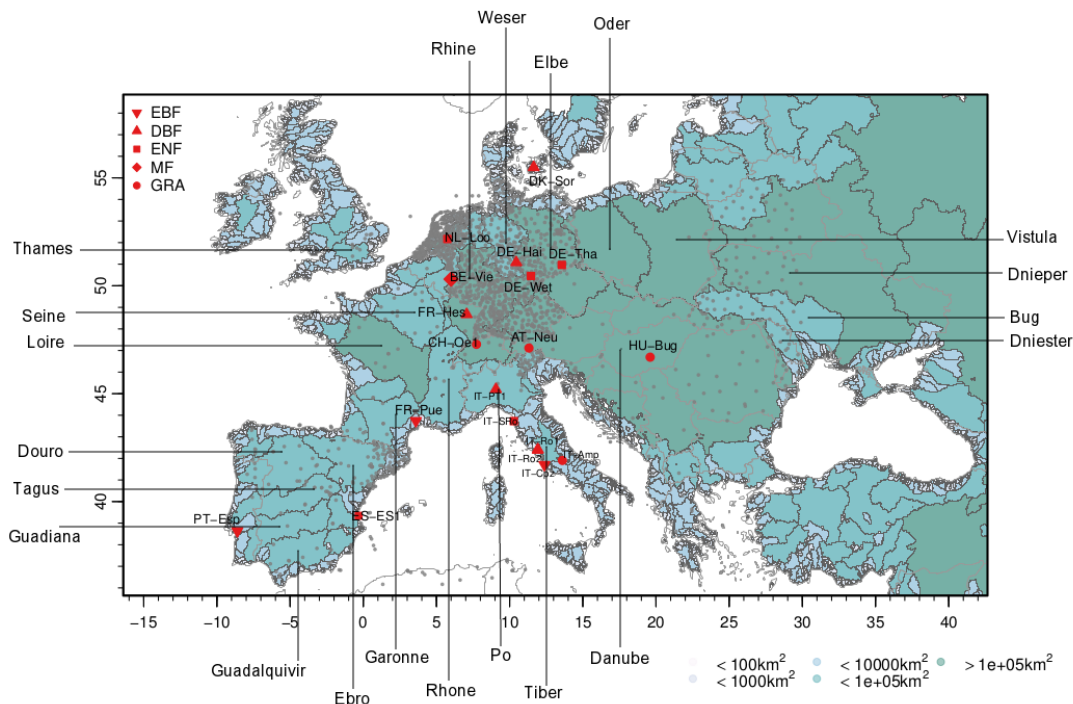
Figure 5.1: European watersheds larger than 100 km^2 considered in this study. Colors are representative of the watershed area. Watershed boundaries are taken from the Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS[1], Lehner et al. 2006). Grey and red points indicate *in-situ* observations from ECA&D and FLUXNET stations, respectively. The symbols are commensurate to the dominant land use of the FLUXNET station: evergreen broadleaf forests (EBF), deciduous broadleaf forests (DBF), evergreen needleleaf forests (ENF), mixed forest (MF) and grasslands (GRA).

these observations into 15 observational data sets of $ET - P$ to assess the uncertainty of the observed net moisture flux at the watershed scale.

Second, co-located *in-situ* measurements of ET and P from FLUXNET eddy-covariance sites are used to validate local-scale and daily variations of the net moisture flux, and to identify the added value of incorporating HWU in fully-integrated modeling systems at the local scale. In addition, a larger data set of *in-situ* rain gauge measurements is used to identify skills at the local scale. Tables 5.1 and 5.2 provide an overview of the gridded and *in-situ* observations used in this study, their provenance and spatial and temporal resolution. A detailed description of each data set is provided below.

Gridded observations

CRU   CRU TS4.00 (Harris et al., 2014) is a gridded time series data set, which was released in January 2017. It provides global, monthly observations of P at 0.5° resolution. Monthly station measurements are interpolated to a high-resolution grid by the Climate Research Unit (CRU) of the University of East Anglia (`http://www.cru.uea.ac.uk/`) to provide a global, homogenous climatic time series of basic variables.

GPCC   The global P data set from the Global Precipitation Climatology Centre (GPCC) is based on 67200 quality-controlled gauge measurements and contains the monthly records

Name	Description	Time scale	Spatial scale	Variable [unit]	References
PRECIPITATION					
CRU v. 401	interpolated gauge observations, not bias corrected (angular-distance weighting (ADW) interpolation)	monthly	0.5°	P [$mm/month$]	Harris et al. (2014)
GPCC	interpolated gauge observations, not bias corrected	monthly	0.5°	P [$mm/month$]	Rudolf et al. (2005)
UDel P	interpolated gauge observations	monthly	0.5°	P [$cm/month$]	Legates and Willmott (1990)
PREC/L v.4.01 (NOAA)	interpolated gauge observations, not bias corrected	monthly	0.5°	P [mm/day]	Chen et al. (2002)
MSWEP v.1.2	multi-source weighted ensemble observations (satellites, gauges and reanalysis)	monthly (aggregated from 3h)	0.1°	P [mm/day]	Beck et al. (2017a, 2017b)
EVAPOTRANSPIRATION					
FLUXNET-MTE	Statistical upscaling of flux observations using a Model-Tree-Ensemble Approach	monthly	0.5°	LE [$MJ/m^2/d$]	Jung et al. (2009, 2010, 2011)
GLEAM v.3.1a	Priestley and Taylor based algorithm based on satellite observations	monthly (aggregated from daily)	0.25°	ET [mm/day]	Miralles et al. (2011), Martens et al. (2017)
CSIRO	derived from the Budyko framework with observed P and Penman-Monteith E	monthly	0.5°	ET [$mm/month$]	Leuning et al. (2008), Zhang et al (2016)

Table 5.1: Overview of the gridded observational data sets of P and ET used in this study.

of P interpolated to a regular 0.5° x 0.5° grid (Rudolf et al., 2005; Schneider et al., 2011). This gridded gauge-analysis product is not bias corrected and may include systematic gauge errors.

UDelP The P data set from the University of Delaware (UDelP, Willmott and Matsuura, 2001; Legates and Willmott, 1990) combines gauge measurements of P from multiple climate monitoring networks, such as the Global Historical Climatology Network (GHCN2). The

gauge observations were interpolated to a regular 0.5° x 0.5° grid using the Climatologically aided interpolation (CAI, Willmott and Robeson, 1995), and were not corrected for sensor biases.

PREC/L   The National Oceanic and Atmospheric Administration (NOAA) released a reconstructed P product over land (PREC/L) based on 17000 gauge observations from the monitoring networks GHCN2 and Climate Anomaly Monitoring System (CAMS). Observations were interpolated to monthly averaged P over a regular 0.5° x 0.5° grid (Chen et al., 2002).

MSWEP   The multi-source weighted ensemble precipitation (MSWEP) is a new P product, which was released in 2017 (Beck et al., 2017a). It provides 3-hourly P at 0.1° resolution worldwide and combines gauge, satellite and reanalysis data in a weighted ensemble approach. MSWEP adjusts for gauge biases on a daily basis and has been developed specifically for hydrological applications (Beck et al., 2017b).

FLUXNET-MTE   The model tree ensemble (MTE) proposed by Jung et al. (2009, 2010, 2011) is a machine learning approach to upscale eddy covariance measurements at FLUXNET stations (described below) to a global gridded data set. The machine learning technique was trained using remote sensing data sets, land use information, climate and meteorological data (Jung et al., 2011). Thus, in regions of sparse FLUXNET measurements, the resulting fields rely strongly on the information content and accuracy of the training data set. The FLUXNET-MTE data set provides monthly averages of e.g., ET at 0.5° resolution, and is often used as an independent data set for model evaluation and calibration (e.g., Swenson and Lawrence, 2014; Bonan et al., 2011).

GLEAM   The Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011) estimates ET based on satellite observations. In GLEAM, ET is calculated as the residual of potential ET from the Priestley-Taylor equation constrained by water stress obtained from satellite-observations of soil moisture and vegetation optical depth, and conditioned on estimates and reanalyses of P, radiation and temperature. In this study, the GLEAM version 3.1a was used (Martens et al., 2017). GLEAM provides daily estimates of ET at 0.25° resolution, which have been shown to agree well with existing FLUXNET stations (Martens et al., 2017). GLEAM retrievals of root zone soil moisture and ET are often used in land-atmosphere feedback studies (e.g., Guillod et al., 2015, Miralles et al., 2012).

CSIRO   The ET data set from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Australia provides global estimates of land surface ET and its components at 0.5° resolution. ET is derived in a Budyko framework using observed P and potential ET calculated with the Penman-Monteith-Leuning model (Leuning et al., 2008; Zhang et al., 2016).

Name	Description	Time scale	Spatial scale	Variable [unit]	References
FLUXNET	Eddy-covariance measurements	daily (aggregated from 1/2 hourly)	*in-situ*	P [mm/day], LE [W/m^2]	Baldocchi et al. (2001)
ECA&D	weather stations/ rain gauges	daily	*in-situ*	P [mm/day]	Klein Tank et al. (2002)

Table 5.2: Overview of the *in-situ* observational data sets of P and ET used in this study.

In-situ observations

FLUXNET The European FLUXNET community (e.g., Baldocchi et al., 2001; `http://fluxnet.fluxdata.org/`) provides an error-corrected and quality-checked, combined data set of measurements across registered European eddy-covariance stations with half-hourly resolution. For 2003, a total of 19 stations over Central and Southern Europe are considered for validation, using a data coverage of at least 50% for P and ET, respectively. This selection comprises 4 eddy-covariance stations over grasslands (GRA), 5 stations over evergreen needleleaf forests (ENF), 6 stations over deciduous broadleaf forests (DBF), 3 stations over evergreen broadleaf forests (EBF) and 1 station over mixed forest (MF). 3 stations (i.e., IT-Col (DBF), NL-HAA (GRA) and CZ-BK1 (ENF)) were neglected for analysis due to quality concerns. Figure 5.1 provides an overview of the FLUXNET stations used for validation. Note that uncorrected measurements of the latent heat flux are used for comparison, and that eddy covariance stations do not necessarily close the land surface energy budget (e.g., Wilson et al., 2002).

ECA&D The European Climate Assessment & Data (ECA&D; Klein Tank et al., 2002; `http://www.ecad.eu/`) collection provides a comprehensive data set of daily in situ observations of multiple variables including P. Figure 5.1 shows the spatial distribution of P stations available for 2003 across Europe. Only stations with at least 95% data coverage (i.e., 342 days of 360 days excluding the model spinup) were selected for validation. This yields a total of 1033 stations across Europe.

Data aggregation

For all gridded observations at different spatial resolutions, watershed means are calculated based on the original resolution (native grid) and compared to the watershed average of the simulations. Daily values of gridded data sets, i.e., GLEAM and MSWEP, are summarized to monthly values. The half-hourly values of FLUXNET are accumulated to daily latent heat fluxes and converted to daily ET (excluding missing values consistently from model results) and daily P, respectively. For the validation with station data, simulations are interpolated to the stations using the nearest neighbor method. For the comparison with

FLUXNET stations, the nearest neighbor is used irrespective of the predominant land use type.

The following relationships are used to convert latent heat fluxes LE in $[W/m^2]$ to ET in $[mm/hour]$:

$$LE = \frac{3600 \; s}{1 \; h} \frac{1}{\lambda} \cdot ET \tag{5.2}$$

with the latent heat flux of vaporization $\lambda = 2,4501 * 10^6 \; J/kg$. Latent heat flux units are converted using

$$LE = \frac{1}{0,408} \cdot ET \tag{5.3}$$

between LE in $[MJ/m^2/day]$ and ET in $[mm/day]$.

5.2.3   Validation metrics and skill scores

The quality and accuracy of the simulation experiments is evaluated with observations at the monthly, seasonal and annual watershed scale, and at the daily and local scale. The following measures are used. The bias indicates the average direction of the error between simulation f and observation o, i.e.,

$$bias = \frac{1}{n} \sum_{i=1}^{n} (f_i - o_i) \tag{5.4}$$

over a time series of length n. f and o are ET, P or their difference from simulation and observation, respectively. The bias provides only an average error between simulation and observation and does not indicate the quality of the simulations. The mean squared error (MSE) is used to measure the accuracy of the forecast, i.e., the magnitude of the error weighted with the squares of error as

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (f_i - o_i)^2. \tag{5.5}$$

The $MSE\ skill\ score$ (SS_{MSE}; Jolliffe and Stephenson, 2012) relates the improvement of one simulation over another, and allows to identify potential improvements of accuracy between the water use scenarios and the natural reference simulation. The SS_{MSE} is calculated as

$$SS_{MSE} = \frac{MSE_{HWU} - MSE_{NAT}}{MSE_{perf} - MSE_{NAT}} = 1 - \frac{MSE_{HWU}}{MSE_{NAT}} \tag{5.6}$$

where MSE_{HWU} is the mean squared error of each water use scenario and MSE_{NAT} is the mean squared error of the natural reference simulation. MSE_{perf} describes the error of a perfect forecast with $MSE_{perf} = 0$. Values close to 0 indicate no change of accuracy by incorporating HWU. A value of 1 indicates an improvement of 100%. The SS_{MSE} is a positively oriented score, i.e., the higher the value, the stronger the improvement.

Similarly to Kotlarski et al. (2017), we consider observational uncertainty in our comparison. Here, observational uncertainty is defined as the standard deviation of the variable's values among the observations.

Daily, *in-situ* observations also allow to evaluate deterministic forecasts in a probabilistic approach, based on, e.g., the joint occurrence of events. For P, we define thresholds for wet days (WD, $P > 1\ mm/day$), heavy precipitation days (HPD, $P > 10\ mm/day$) and very heavy precipitation days (VHPD, $P > 20\ mm/day$), thus converting continuous values of P into dichotomous data sets of *yes/no* events. These events are then evaluated probabilistically by their joint occurrence in simulations and observations. Therefore, a contingency table (Tab. 5.3) is defined for each threshold, which allows to evaluate the conditional probability of an event being simulated, given the observations, and vice versa.

		OBSERVATION	
		yes	**no**
SIMULATION	**yes**	e_h hits	e_{fa} false alarms
	no	e_m misses	e_{cn} correct negatives

Table 5.3: Sample 2x2 contingency table for dichotomous (yes/no) events between simulation and observation. The notations show (e_h) hits, (e_{fa}) false alarms, (e_m) misses and (e_{cn}) correct negatives.

A number of skill scores have been identified based on the contingency table (e.g., Jolliffe and Stephenson, 2012), such as the frequency bias (FBI)

$$FBI = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}}\ , \tag{5.7}$$

which indicates the tendency of the simulations to over- or underpredict events. A value larger than 1 indicates that more events were simulated than observed. Vice versa, values below 1 indicate underprediction. The Peirce's skill score (PSS) evaluates how well the simulations are able to separate the occurrence of events (*yes* events) from the non-occurrence (*no* events). The PSS is defined as

$$PSS = \frac{\text{hits}}{\text{hits} + \text{misses}} - \frac{\text{false alarms}}{\text{false alarms} + \text{correct negatives}}\ , \tag{5.8}$$

and hence evaluates the relation between probability of detection (first term) and the probability of false detection (second term). The PSS varies between -1 and 1, where 0 indicates no skill. A perfect separation between occurrence and non-occurrence of events is indicated by a PSS of 1.

5.3 Validation of the natural reference simulation

This section evaluates the accuracy of the natural reference simulation with TerrSysMP over the heatwave year 2003. It is split in two parts: (i) the monthly watershed scale, at which observational uncertainty of gridded products is considered, and (ii) the daily, local scale using *in-situ* observations.
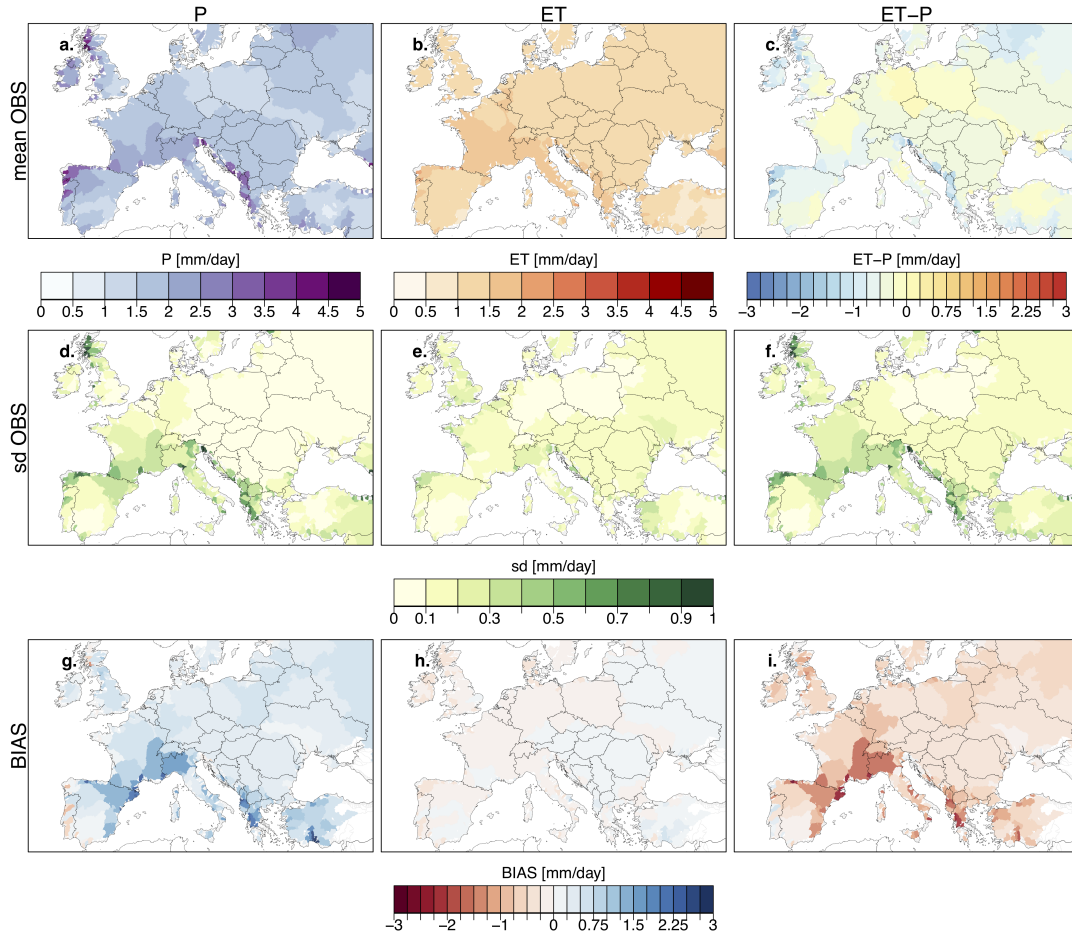
Figure 5.2: Maps of annually averaged P $[mm/day]$, ET $[mm/day]$ and $ET - P$ $[mm/day]$ over European watersheds in 2003. (a-c) and (d-f) show the mean and the standard deviation over the 5, 3 and 15 observations (each combination of ET: FLUXNET-MTE, CSIRO, GLEAM with P: GPCC, CRU, PREC/L, MSWEP, UDelP), respectively. (g-i) shows the mean bias of the NAT simulation of TerrSysMP with respect to the mean observations, again for P, ET and $ET - P$.

5.3.1   Watershed scale

In order to consider the observational uncertainty, we evaluate the accuracy of the simulated terrestrial water cycle based on multiple gridded observational products of ET and P at the watershed scale. Figure 5.2 shows ET, P and $ET - P$ averaged over the watersheds for the year 2003. The upper two rows show the average and the standard deviation over the 5, 3 and 15 observational products, respectively. The bottom row shows the bias of the natural reference simulation with respect to the mean of all observations. Over the year, the average P commonly exceeds ET, but a few watersheds show a positive net moisture flux (e.g., Guadalquivir, Guadiana, Seine, Loire, Elbe and Oder; Fig. 5.2c). The model shows a positive P bias, especially over mountainous watersheds. The ET bias is comparably small, which leads to an underestimation of the net moisture flux by the model over almost all watersheds (Fig. 5.2i). Note, the highest observational uncertainties, arising through P, coincide with the highest biases (Fig. 5.2d+g; Fig. 5.2f+i). Especially in mountainous regions, such as the Alps, the Pyrenees, the Cantabrian mountains (NE Spain) and the
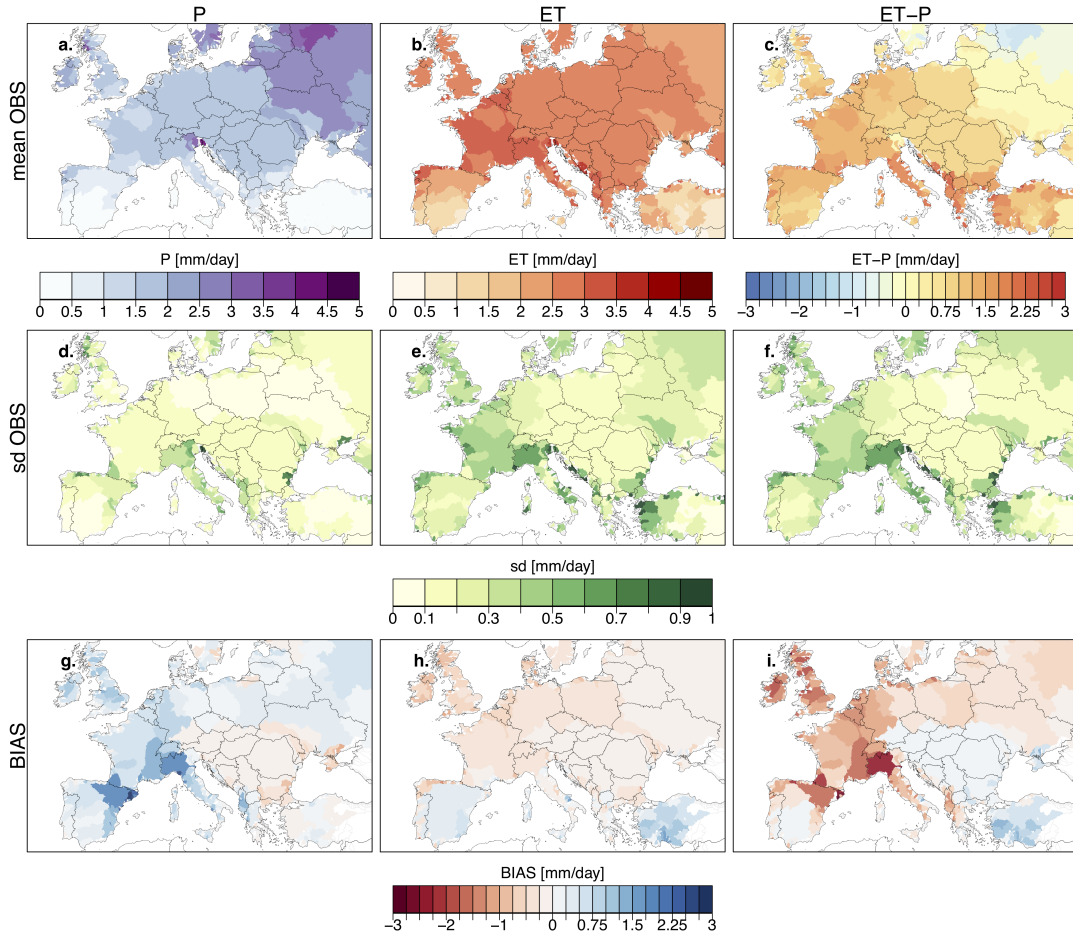
Figure 5.3: Same as Fig. 5.2 but for summer 2003 (June-July-August).

mountain ranges of the Iberian system (NW Spain), $ET - P$ is underestimated, but the observations exhibit a high uncertainty.

During the dry summer (June, July, August) of 2003, all European watersheds lost more water through ET than they received through P, which resulted in a positive upwards moisture flux (Fig. 5.3a-c). Most watersheds over France, the Iberian Peninsula and the Mediterranean received almost no P and indicate a net loss of water on the order of $2 - 3mm/day$. Whereas the model captures the net loss of water over the Southern European watersheds quite well, it simulates a negative $ET - P$ balance over the watersheds Ebro, Rhone and Po, resulting in a strong negative bias (Fig. 5.3i), which also indicates the lack of skill in simulating heatwaves (Weisheimer et al., 2011). Again, this coincides with the highest observational uncertainty, determined by observational products of ET during summer (Fig. 5.3e+f). Interestingly, the Po watershed, which is heavily water managed (Wada et al., 2014), shows the largest observational uncertainty of $ET - P$ during summer.

Observational uncertainty is mainly determined by P over the course of the year, and by ET during summer. Figure 5.4 shows exemplary the annual cycle of ET, P and $ET - P$ for the Guadalquivir basin. The annual cycle of the water fluxes is well captured for this watershed, but also indicates some differences between the model and the observations. The observations show a peak of ET in late spring (Fig. 5.4a), presumably through the

prolonged P deficit and the pre-heatwave from March to May. However, the ET observations also show a large uncertainty from April to September, which is mainly stemming from GLEAM, indicating the strongest water-limitation of the observational data sets from June onwards. In contrast, this effect is hardly visible in case of CSIRO and FLUXNET-MTE, which agree better with the simulation results. The origin of the existing biases cannot be traced back conclusively to P, which is only marginally overestimated in this case. Vice versa, it remains unclear to which extent the ET observations are able to capture small-scale irrigation effects. In case of GLEAM, this is limited to the availability of satellite measurements every few days, during which the effect of local scale irrigation might not be visible. Similarly, FLUXNET-MTE is limited to the availability of tower measurements, which might not be located in irrigated areas. Considering all these uncertainties, the annual cycle of ET-P is well captured, with an underestimation in spring and a shift of the peak from May to July.
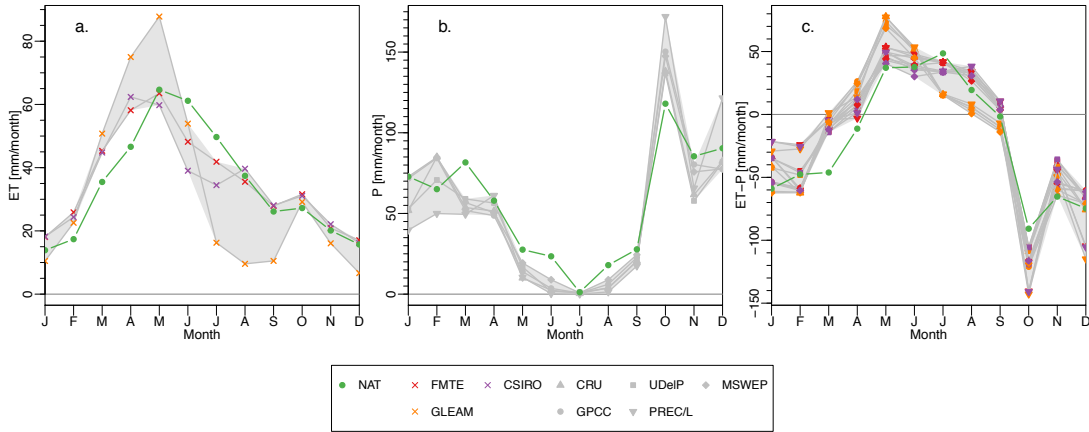
Figure 5.4: Monthly time series of (a) ET, (b) P and (c) $ET-P$ in [mm/month] averaged over the Guadalquivir watershed for all gridded observational data sets and the NAT simulation. The grey shaded areas indicate the maximum range covered by the observations.

5.3.2   Local scale

The validation of daily ET, P and $ET - P$ fluxes with *in-situ* measurements of 1033 rain gauges and 19 FLUXNET towers confirms the above watershed-based biases. Figure 5.5 shows the annual (a-c) and summer (d-f) biases of P, ET and $ET - P$ for all available stations in 2003. Over the year, the simulation exhibits a positive P bias on the order of 0.5 mm/day based on FLUXNET stations, and 0.49 mm/day based on ECA&D stations. This value is comparable to other validation studies (e.g., Kotlarski et al., 2017; Kotlarski et al., 2014; Vautard et al., 2013; Katragkou et al., 2015). In contrast to the watershed scale, ET is overestimated, with a bias of 0.24 mm/day. For all available co-located measurements, this results in an underestimation of $ET - P$ on the order of -0.29 mm/day. These biases are slightly increased for summer (0.94 mm/day (P, FLUXNET), 0.69mm/day (P, ECA&D), 0.46 mm/day (ET, FLUXNET), -0.48 mm/day ($ET - P$, FLUXNET)). Noteworthy is the high P bias over all stations in the area of Ebro and the Alps, which corroborates the well-known dependence on topography. On the other hand, the biases at

FLUXNET stations reveal no dependence on the land use, which suggests the dominant role of P biases in the simulation. Yet, it has to be mentioned that the lack of stations, and especially co-located stations, is a major limitation for this and many other studies.

A detailed analysis of the daily P statistics reveals that heavy and very heavy P events are most commonly responsible for an increased P bias. Figure 5.6 shows boxplots of the FBI and the PSS for WD, HPD and VHPD, separately for the regions Mid-Europe (ME), France (FR), Iberian Peninsula (IB), Mediterranean (MD) and Eastern Europe (EA). The FBI shows only a small overprediction of WD, with the highest skill in ME and FR. The FBI increases with the P threshold and shows an overprediction of HPD and VHPD. Over MD and IB, HPD (VHPD) are on average 2-3 (6-8) times more often simulated than observed. The skill also continuously decreases, and indicates that over the depicted regions the simulation is unable to correctly detect timing and location of VHPD.
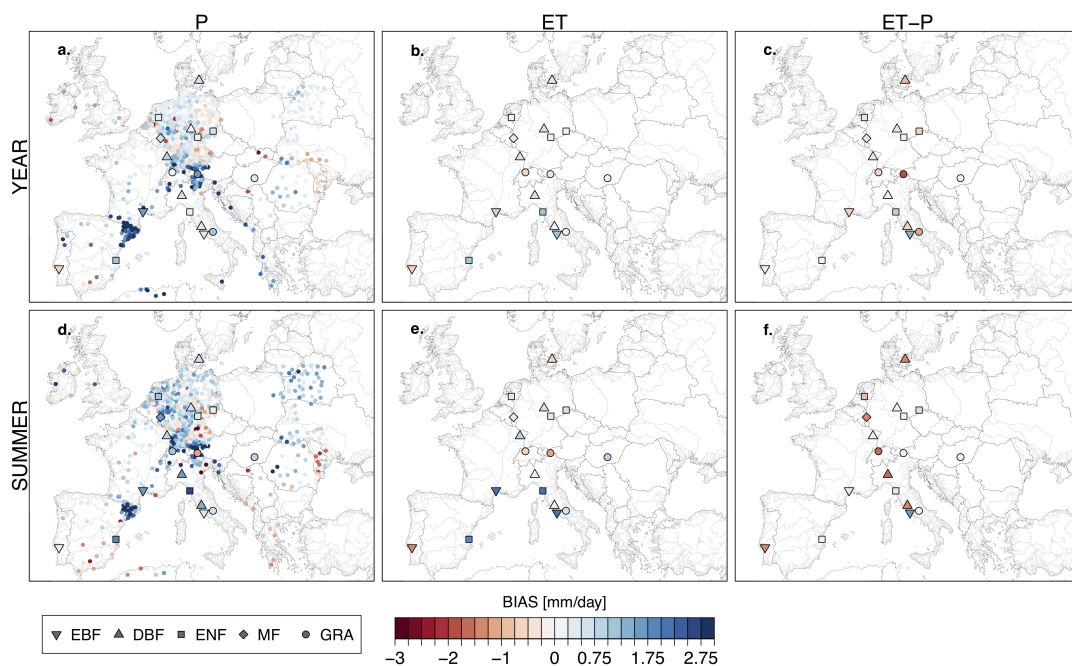
Figure 5.5: Daily P, ET and $ET - P$ biases from the NAT simulations of all available ECA&D and FLUXNET stations for (a-c) the full year 2003, and (d-f) summer 2003. ECA&D stations are illustrated by colored dots. FLUXNET stations are indicated by a black border around the symbols, which indicate the land use of the station.
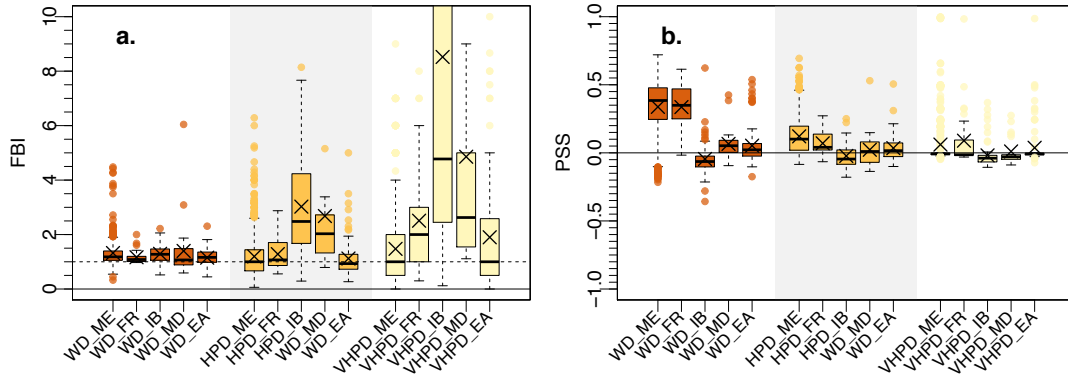
Figure 5.6: (a) Frequency bias (FBI) and (b) Peirce's skill score (PSS) of the NAT simulation for wet days (WD, brown), heavy precipitation days (HPD, orange), and very heavy precipitation days (VHPD, yellow) over the full year 2003. The boxplots show the distribution of each score over 696, 30, 76, 25 and 104 stations in the regions Mid-Europe (ME), France (FR), Iberian Peninsula (IB), Mediterranean (MD) and Eastern Europe (EA), respectively. The crosses denote the mean score over these stations.

5.4   Added value of incorporating HWU

The question arises, to which extent the incorporation of HWU increases the skill of the simulations. Here, the added value of incorporating HWU on (i) monthly fluxes at the watershed scale, and (ii) daily, local fluxes of ET, P and $ET - P$ is assessed.

5.4.1   Watershed scale

Figure 5.7a shows the annual bias of $ET - P$ over 15 large watersheds across Europe as a boxplot for each simulation including all observational products. A comparison of the boxplots in each column (i.e., for each watershed) shows that the incorporation of HWU does not remove the overall bias, which can also be seen in Fig. 5.4. The bias is rather subject to observational uncertainty, which increases from flat and coastal areas (Guadalquivir, Guadiana and Tagus) towards mountainous watersheds (Douro, Ebro, Garonne and Rhone). Here, the observational uncertainty clearly exceeds the uncertainty related to HWU.

The high observational uncertainty of $ET - P$ leads to a high uncertainty of the bias especially in summer (Fig. 5.7b). As has been discussed before, the models fail to capture the prolonged drought and exhibit consistently a small overestimation of the net moisture flux especially for the southern watersheds over the Iberian Peninsula. For most of the analyzed watersheds, the bias is negative in summer, and hence indicates an underestimation of $ET - P$, either through an underestimation of ET or an overestimation of P (cf. Fig. 5.3). The summer bias is largest for the CSIRO data sets, ranging from, e.g., -0.9 mm/day to -0.4 mm/day for the Po basin. On the contrary, the biases for Douro, Tiber, Garonne and Elbe can be either negative, or approximately 0, and depend on the observational products used. The fact that the highest biases coincide with the highest observational uncertainty, and furthermore exceed the variations of $ET - P$ induced by HWU, hamper an objective validation. Moreover, this raises concerns especially for the use of gridded validation data sets in water assessment and scarcity studies.
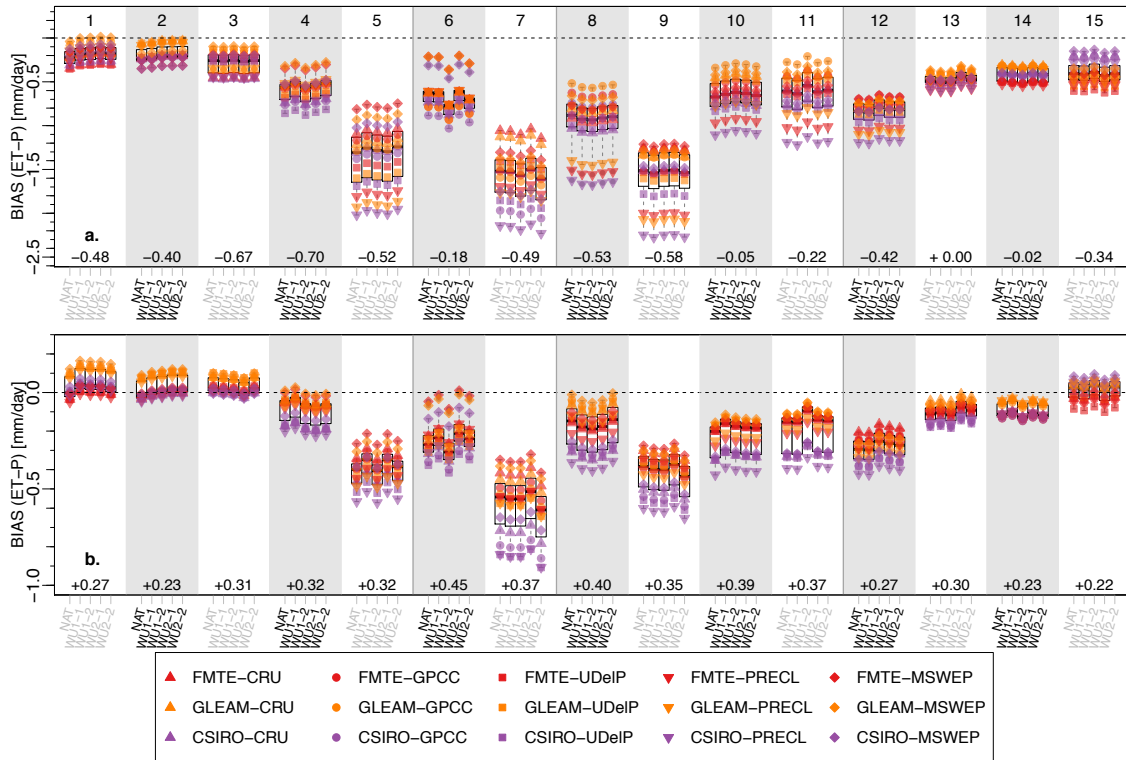
Figure 5.7: Box-Whisker-Plots of the $ET - P$ bias for (a) the full year 2003 and (b) summer 2003 for the 5 simulations, the 15 observations and over 15 selected watersheds: 1-Guadalquivir, 2-Guadiana, 3-Tagus, 4-Douro, 5-Ebro, 6-Tiber, 7-Po, 8-Garonne, 9-Rhone, 10-Seine, 11-Loire, 12-Rhine, 13-Elbe, 14-Oder, 15-Danube. The 5 boxplots for each watershed indicate the observational uncertainty of the simulated bias in the following order: NAT, HWU1-1, HWU1-2, HWU2-1, HWU2-2. The symbols within each box indicate the observational P data set ($\triangle$: CRU, $\bigcirc$: GPCC; $\square$: UDelP, $\triangledown$: PREC/L, $\diamondsuit$: MSWEP) and the color the observational ET data set (red: FLUXNET-MTE, orange: GLEAM, purple: CSIRO). A negative bias in (a) indicates that the net moisture flux $ET - P$ is overestimated (simulations too wet) and in (b) that the continental source is underestimated. A positive bias in (b) indicates that $ET - P$ is overestimated. The values at the bottom of each subplot indicate the mean $ET - P$ flux of all observations over the watershed.

The individual HWU scenarios were not constructed to represented real-world conditions, but rather to assess local and non-local feedback processes in the terrestrial water and energy cycles associated with HWU. Hence, we do not expect the single HWU scenarios to improve model skill with respect to NAT. The mean of all scenarios, HWU-ENS, is assumed to be representative of the effect of a more realistic representation of HWU. Hence, we focus on the added value of HWU-ENS over NAT in the following.

Figure 5.8 shows the SS_{MSE} of HWU-ENS over NAT and its dependency on observational uncertainty. The boxplots show the SS_{MSE} of P, ET and $ET - P$ for the 15 watersheds from Fig. 5.7, for the full year (Fig. 5.8a) and summer (Fig. 5.8b), and include the 5, 3 and 15 observational products, respectively. Most watersheds indicate a clear improvement in the simulation of P and $ET - P$ (e.g., Guadiana, Ebro, Tiber, Elbe), but a deterioration of skill in the simulation of ET over the summer and the full year. The magnitude of this skill improvement for $ET - P$ varies between 1 and 30% (average: 3% for the full year, 4% for the summer). However, for some watersheds the added value is
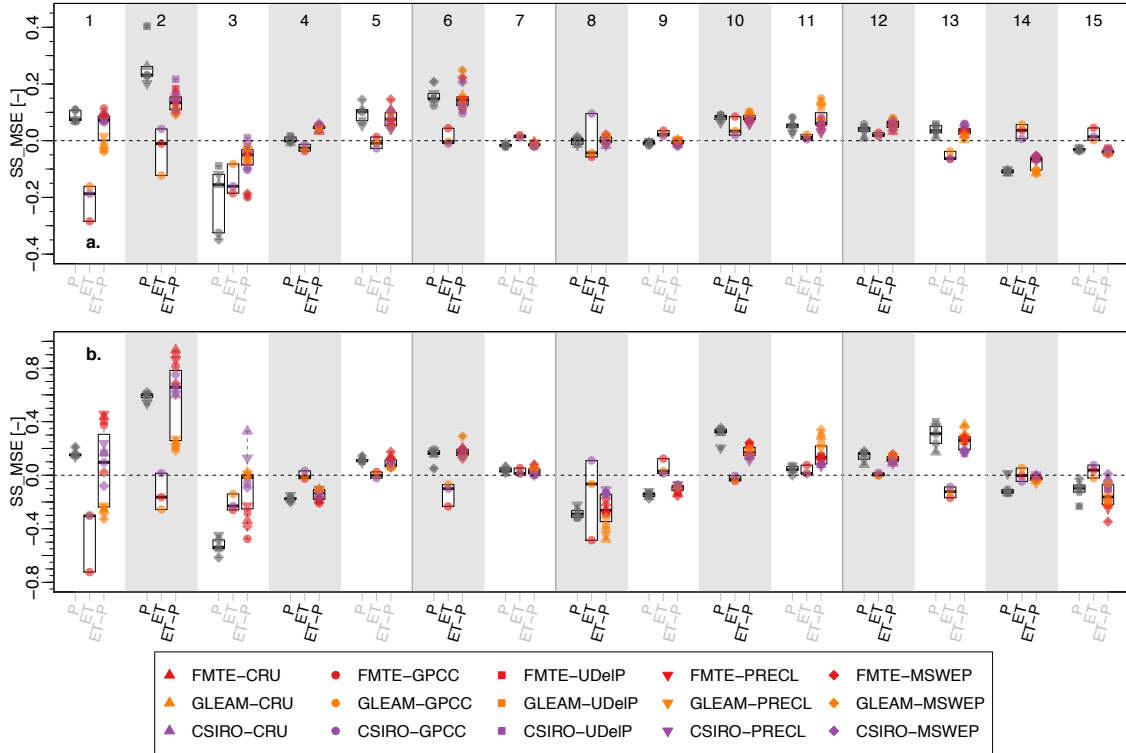
Figure 5.8: Box-Whisker-Plots of the SS_{MSE} for P, ET and $ET - P$ of HWU-ENS over HWU-NAT, for (a) the full year 2003 and (b) summer 2003 over the 15 selected watersheds from Fig. 5.7: 1-Guadalquivir, 2-Guadiana, 3-Tagus, 4-Douro, 5-Ebro, 6-Tiber, 7-Po, 8-Garonne, 9-Rhone, 10-Seine, 11-Loire, 12-Rhine, 13-Elbe, 14-Oder, 15-Danube. The 3 boxplots for each watershed indicate the uncertainty of the SS_{MSE} with respect to the observational data set: P ($\triangle$: CRU, $\bigcirc$: GPCC; $\square$: UDelP, $\triangledown$: PREC/L, $\diamondsuit$: MSWEP), ET (red: FLUXNET-MTE, orange: GLEAM, purple: CSIRO) and $ET - P$. A positive (negative) SS_{MSE} indicates that the incorporation of HWU increases (decreases) the skill of the simulations. A value of 0 indicates no change of skill.

masked by observational uncertainty, either in magnitude (e.g., Guadiana) or in sign (e.g., Guadalquivir and Garonne). Over all watersheds (including the ones not illustrated here), the SS_{MSE} indicates the incorporation of HWU-ENS improves marginally the simulation of ET by 0.3%, but deteriorates the skill for P and $ET - P$ fluxes by some 1%, which is insignificant against the background of observational uncertainty. On the contrary, the simulation of the summer net moisture flux is improved by about 8%, which mainly stems from an improved simulation of P (11%).

5.4.2   Local scale

The differences associated with HWU increase toward the daily and local scales (Keune et al., 2018). To assess the added value of incorporating HWU at the local scale, daily *in-situ* observations of ET, P and $ET - P$ are used. A SS_{MSE} is calculated for the 19 available FLUXNET stations. The results are shown in Fig. 5.9, for $ET - P$, ET and P along the rows, and for the entire year and summer along the columns, respectively. Over the year and all available FLUXNET stations, the HWU mean improved the MSE of $ET - P$ by 13%, of P by 13% and ET by 3%. The average improvements are larger in summer (ET-P: 27%, P:

27%, ET: 4%). This improvement can be significantly higher if individual HWU scenarios are inspected. However, effects cannot be directly associated with individual PFTs, regions or HWU scenarios (cf. Fig. E.3). The separation of the skill score into ET and P shows that the largest increase in skill is coming from an improved simulation of P, although especially summer skills seem arbitrarily affected due to the chaotic nature of the P. Note, while summer ET at PT-Esp is significantly improved through the incorporation of HWU for all water use scenarios (approx. 10 to 40%), the skill of $ET - P$ does not improve, as P skills are not improved (cf. also Fig. E.4).
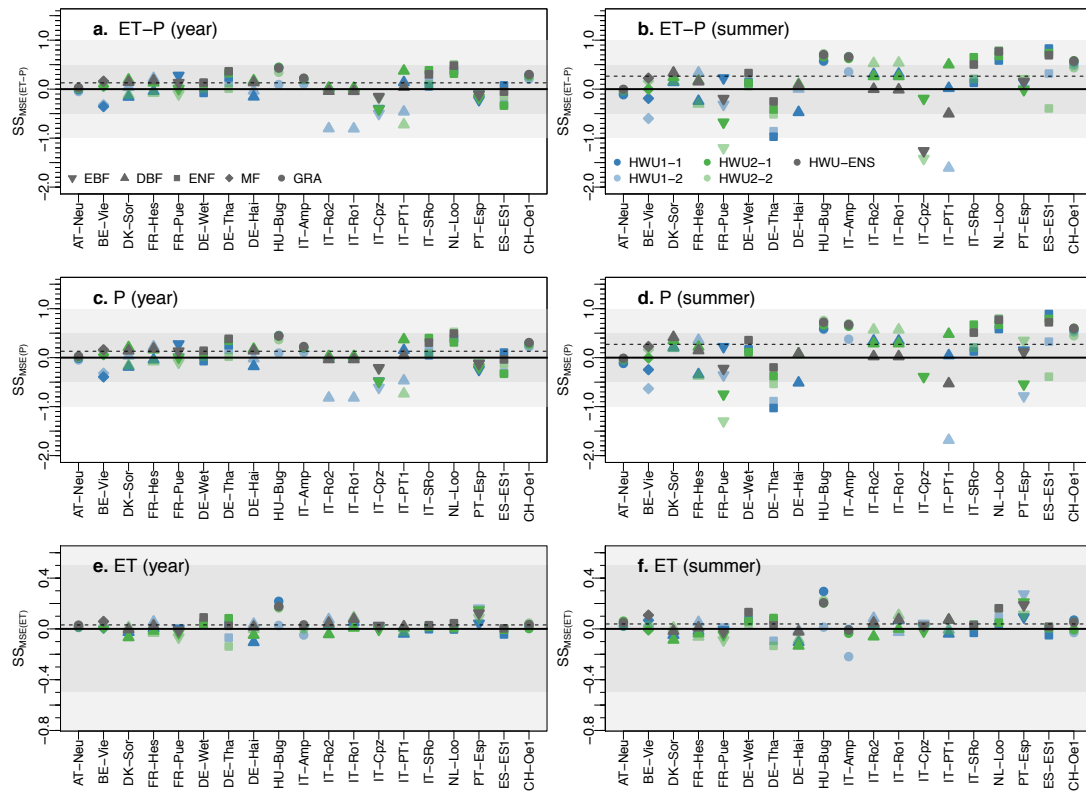
Figure 5.9: SS_{MSE} for improvements of the water use scenarios over the natural reference run for daily $ET - P$, ET and P for the full year and summer 2003, respectively. The symbols indicate the dominant land use of the FLUXNET tower. The grey points indicate the average improvements of the water use ensemble mean over the reference run for all stations. The dashed black line shows the average improvement of the water use ensemble over the reference simulation.

As the evaluation at FLUXNET sites indicates major improvements for the simulation of P, individual events are further evaluated with a larger observational data set, to identify an overall added value. Figure 5.10 shows spatial maps of the SS_{MSE} for daily summer P for HWU-ENS and the individual HWU scenarios. Spatially, the skill improvement varies, and simulated P at single stations is not consistently improved through the incorporation of HWU. Moreover, all stations, including those in less managed regions, indicate changes in skill, which emphasizes the remote impacts of HWU. However, there exist some common patterns of skill improvement for the individual HWU scenarios. The majority of stations over the Iberian Peninsula show a negative influence of HWU on the simulation of daily summer P for all HWU scenarios (on average by -9%, cf. Fig. E.5). In contrast, the
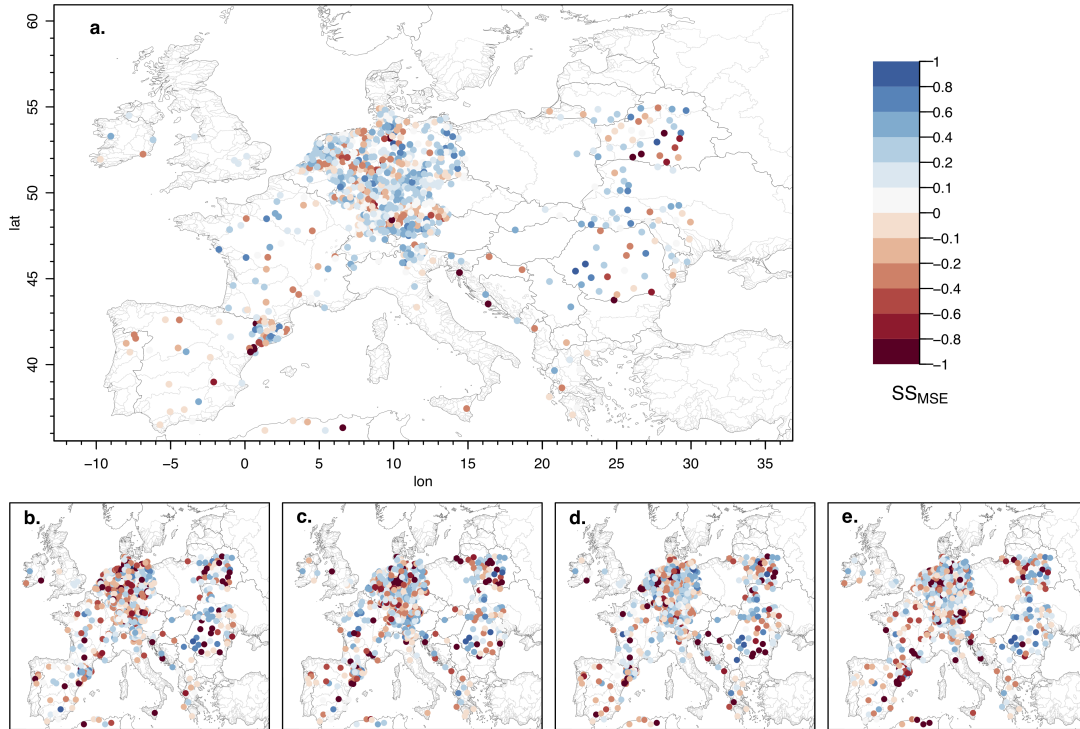
Figure 5.10: Map of SS_{MSE} for (a) the HWU ensemble mean, (b) HWU1-1, (c) HWU1-2, (d) HWU2-1 and (e) HWU2-2 relative to the NAT simulation for summer 2003. Blue (red) dots indicate an improved (deteriorated) skill in simulating P through the incorporation of HWU.

majority of the stations over Germany and Eastern Europe, with a focus on the Eastern parts of the Danube watershed, indicate improved skill (on average by 12.9% over ME). Yet, while the skill of individual HWU scenarios is affected by varying skill through over- and underestimated P and the incorporated uncertainty of HWU (full year: HWU1-1 (-0.5%), HWU1-2 (-0.7%), HWU2-1 (-0.5%), HWU2-2 (-0.6%); summer: HWU1-1 (3.1%), HWU1-2 (2.9%), HWU2-1 (6.7%), HWU2-2 (-2.7%)), the skill improvement of the HWU-ENS indicates an increase of information contained in the simulations through the incorporation of HWU by +7.9% for the full year and 18.96% for summer.

The associated FBI and PSS are very similar between HWU-ENS, the individual HWU scenarios and NAT (Fig. E.6), but indicate that the incorporation of HWU potentially reduces the FBI for VHPD over ME (from 1.92 to 1.61) FR (from 2.16 to 1.84), MD (from 5.84 to 5.25) and EA (from 1.90 to 1.58), but increases over IB (from 8.52 to 9.00). However, overall, large uncertainties remain in attributing local skill improvement, which is potentially related to use of a parameterization scheme for convection and the lack of stations over large parts of Southern Europe.

5.5 Discussion

5.5.1 Observational uncertainty

This study evaluated the uncertainty of bias estimates from simulations using multiple observational data sets, from the local to the watershed scale. The obtained biases of approximately 0.5 mm/day for P and 0.24 mm/day for ET are comparable to other studies (e.g., Mueller and Seneviratne, 2014; Kotlarski et al., 2017; Kotlarski et al., 2014; Vautard et al., 2013; Katragkou et al., 2015), and lead to a bias of the net water flux $ET - P$ on the order of -0.29 mm/day. While the incorporation of human water use apparently does not remove these wet biases, this result is subject to observational uncertainty. The influence of observational uncertainty in ET and P on the net water flux bias varies, and indicates reasonable limitations that, at least partly, explain the differences arising from observational products through, e.g., limitations of remotely sensed products. While this uncertainty has previously been identified (e.g., Mueller et al., 2013; Mueller et al., 2011), the usefulness of these products to validate and potentially calibrate models has not been discussed yet. The challenges and risks of validating hydrologic- and land surface models with one gridded observational product of e.g., ET are obvious, as *in-situ* measurements remain sparsely located and limited in availability. Based on the results of this study, multiple observational data sets should be used and individual simulations should be regarded as single scenarios of a multi-model and multi-observational ensemble, in order to assess uncertainty of water availability under current and future conditions. Note that the true error remains unknown, because the true fluxes are not known everywhere and every time from direct measurements. Therefore, not only more accurate model simulations are required, but more accurate observational products, which can be used to validate models and allow to assess changes of the terrestrial water cycle exhibiting a smaller uncertainty than the observations.

In addition, the high uncertainty of gridded ET and P observations emphasizes the advantage of coupled model systems, which do not directly rely on external observational data sets, and hence reduce the uncertainty related to e.g., atmospheric forcings (e.g., Biemans et al., 2009). While these models introduce additional biases with respect to, e.g., precipitation, which may propagate from one compartment to the other, internal feedback processes and storages are physically consistent. Thus, the origin of biases from observations or simulations cannot be easily identified and constitutes a major challenge for validating coupled modeling systems (Bauer et al., 2015). For future studies, we propose the use of probabilistic approaches to tackle these issues, and the incorporation of observational uncertainty estimates in validation utilizing extreme ensemble approaches with a very large number of ensemble members to simulate the terrestrial water cycle.

5.5.2 Potential for added value of incorporating HWU

Furthermore, this study assessed the added value of incorporating HWU in the simulation of monthly watershed and daily local fluxes of ET and P, as well as $ET - P$. The results do not conclusively show that the incorporation of HWU leads to an added value in the simulation of the terrestrial water cycle, as summarized in Tab. 5.4. Especially at the watershed scale,

	spatial scale	time scale	$SS_{MSE}(P)$	$SS_{MSE}(ET)$	$SS_{MSE}(ET-P)$
year	watershed	monthly	−1.23%	+0.25%	−1.01%
	local	daily	+13.12% +7.85%	+3.17%	+12.82%
summer	watershed	monthly	+11.64%	−0.60%	+8.03%
	local	daily	+27.30% +18.96%	+3.88%	+26.52%

Table 5.4: Summary of MSE-skill-scores evaluated in this study. The skill scores are illustrated for the human water ensemble (HWU-ENS) relative to the natural reference simulation (NAT). The watershed mean relates to the mean of all observations. The two values for the local, daily scale relate to FLUXNET stations (19 stations; top value) and the ECA&D stations (1033 stations; bottom value), respectively.

the incorporation of HWU may lead to a decreased skill of P and $ET-P$ at the annual time scale, and vice versa to a decreased skill of ET in summer. These results do not agree with the daily, local scale analysis, which indicates average improvements for all fluxes for the full year and summer. There are several explanations for this apparent inconsistency. First, the SS_{MSE} for watersheds was calculated using the mean of all observational products, assuming that these represent the true value. Secondly, there is a lack of *in-situ* stations, which measure ET and P simultaneously. Stations are furthermore not equally and only sparsely distributed across Europe, which may also lead to differences from the watershed to the regional scales.

Furthermore, there are two reasons why the use of HWU-ENS leads to significantly better skill scores than the use of the individual HWU scenarios. As has been mentioned before, the HWU scenarios were not constructed to represent real-world conditions, but rather to assess the uncertainty of the associated feedbacks. This might lead to a consistent under- and overestimation of individual scenarios, which is averaged out in HWU-ENS. The averaging in HWU-ENS furthermore leads to a much smoother time series with less extremes, which is potentially causing the increase in skill of simulating VHPDs.

The HWU scenarios did not improve the bias from the NAT simulation. It is noteworthy that differences cannot easily be attributed to local effects of HWU or land use, because we are comparing a 156 km^2 grid box estimate with an *in-situ* measurement over a much smaller footprint based on a nearest neighbor interpolation, irrespective of the land cover. Given the high uncertainty of the gridded observations, and the sparse *in-situ* network, it is hence difficult to conclude that the incorporation of HWU improves the simulation of ET and P. Nevertheless, we find consistent patterns, which indicate the potential of an added value.

5.6   Summary and Conclusions

This study evaluates the potential added value of integrating HWU in simulations of the continental scale integrated Terrestrial Systems Modeling Platform (TerrSysMP) on the simulation of the terrestrial water cycle during the heatwave 2003 in Europe.

In the first part, the reference simulation at the seasonal to annual watershed scale is evaluated, considering observational uncertainty. The study finds that observations of the net moisture flux, $ET - P$, exhibit relatively large spatial uncertainty at the watershed scale, which is highest in mountainous watersheds. TerrSysMP simulations reveal a wet bias at the annual and seasonal watershed scale, which is on the order of 0.2 to 2 mm/day but is strongly dependent on the observational data sets used for validation. The incorporation of HWU does not remove this wet bias in general, however, the observational uncertainty in mountainous watersheds is larger than differences induced by HWU, which precludes the identification of an added value at the watershed scale. While the annual cycle of all fluxes is well reproduced, especially Southern European watersheds exhibit a wet bias, which stems from overestimated P. Watershed averaged ET is especially uncertain during dry and hot summers under predominantly water limited conditions, which are not consistently captured by the observational ET data sets used in this study, due to the small footprint of the eddy-covariance stations and the time and space limitations of satellite ET observations (Lawston et al., 2017). Analogously, it remains unclear to which extent these observations capture the impact of small-scale irrigation. Furthermore, gridded observational P data sets are prone to diverge in regions with a low density of rain gauge observations, which adds to the uncertainty from the different interpolation methods used.

While this study does not conclusively show that the incorporation of HWU leads to an added value in the simulation of the terrestrial water cycle, potential skill improvement for daily ET and P at the local scale is apparent. Unfortunately, co-located measurements of ET and P are scarce, but the 19 FLUXNET stations over Europe indicate an average improvement of daily ET and P values, which is independent from the local land cover. A separation of ET and P shows that the major improvements are in fact arising from an improved simulation of P. A skill assessment with a larger, daily *in-situ* P data set indicates that the incorporation of HWU in an integrated modeling system exhibits potential added value. In case of TerrSysMP, skill is mainly improved in Mid-Europe and France, but deteriorated over the Iberian Peninsula. In fact, P events over the Iberian Peninsula do not show any skill in any of the simulations and show high frequency biases for heavy and very heavy P events.

In general, a more accurate simulation of P is needed to improve the simulation of the terrestrial water cycle. The high uncertainty associated with P suggests the largest potential for improvement. This can either be achieved through bias correction (Rojas et al., 2011), an increasing resolution to circumvent the use of simplified convection parameterizations, and data assimilation. The actual added value might still be larger, if assessed at the meso-β scale (Feser, 2006). Moreover, an improved simulation of ET does not necessarily lead to an improved simulation of P, which is highly dependent on the convection parameterization (Hohenegger et al., 2009). In the future, one could bridge the gap between land surface ET and P by using, e.g., sounding profiles to assess the propagation of improved land surface states into the atmosphere. Ultimately, remote sensing products should be used to validate long-term changes, also in, e.g., subsurface storage (GRACE), to evaluate an intensification of the terrestrial water cycle due to anthropogenic influences and identify the origin of biases in integrated simulations. Here, also longer simulations are needed to assess the long-term

impacts of HWU, as the response of the subsurface is much slower.

Chapter 6

Summary, conclusions and future recommendations

6.1 Summary and conclusions

The growing pressure on water resources via groundwater depletion indicates the need for an improved simulation of the terrestrial system and its feedbacks in order to enable the development of adaptation strategies for a sustainable use of water. Current simulations of terrestrial system processes, however, remain simplified and do not allow for integrated assessments of, e.g., anthropogenic activities, such as groundwater pumping, on the Earth system. This thesis aims to improve the understanding of processes and uncertainties of the terrestrial, continental-scale water cycle related to the groundwater representation and human water use. The objectives of this thesis are to (i) identify and assess impacts of groundwater dynamics on land-atmosphere feedbacks; (ii) evaluate feedbacks of human water use on the continental-scale terrestrial water cycle; and (iii) assess potential added values of incorporating human water use on the simulation of the water cycle. Building upon previous findings, which were mostly restricted to the catchment scale, the first objective focuses on the impact of physics-based groundwater dynamics as a lower boundary condition on land surface and atmospheric states at the continental scale. The second objective focuses on the incorporation of human water use, here defined as groundwater pumping and irrigation, and an evaluation of the combined effects on groundwater and atmosphere. The underlying hypothesis is that groundwater and human water use feedbacks exceed the catchment scale and significantly affect atmospheric processes via groundwater-atmosphere and human water use feedback mechanisms. The third objective describes the identification of an added value arising from the incorporation of human water use in continental-scale, groundwater-to-atmosphere simulations.

In order to address these objectives, the Terrestrial Systems Modeling Platform, Terr-SysMP, is set up over continental Europe and enables simulations of the terrestrial energy and water cycles from the bedrock to the atmosphere, including two-way feedbacks between soil, vegetation and atmospheric processes. Sensitivity studies are performed and allow to assess impacts and uncertainties of the terrestrial systems with respect to the groundwater representation and human water use. The human water use simulations are validated with *in-situ* and remotely sensed observations. Three main research questions, directly related to the aforementioned objectives, are formulated and addressed in three chapters. The answers to these questions are summarized as follows.

> ***How does groundwater contribute to land-atmosphere feedbacks at the continental scale and to which extent might groundwater buffer extreme heat, e.g., during droughts and heatwaves?***

In the first part of this thesis, the impact of a dynamic representation of the groundwater table on land-atmosphere feedbacks is analyzed, using continental-scale groundwater-to-atmosphere simulations over Europe. Therefore, two groundwater configurations, which mimic (i) a typical 1D representation of water table depths in commonly applied land surface schemes, and (ii) a physics-based 3D representation of groundwater dynamics, are set up and simulations are performed over the heatwave in August 2003, during which the largest impact is expected. Results suggest that shallow water tables potentially dampen

temperature extremes through an alleviating soil moisture-temperature feedback. Daily maximum temperatures drop by 1-2 °C in regions of shallow water table depths, while regions with deep water tables are barely affected.

Moreover, groundwater dynamics show significant impacts on land surface, but also atmospheric states, as indicated by, e.g., integrated measures of convective available potential energy (CAPE). However, the extent to which land surface and atmospheric processes are affected, strongly depends on subsurface hydraulic characteristics. Two hydrologic facies distributions (HFDs) representing different types of hydrogeologic heterogeneity are evaluated showing that the prescribed hydraulic conductivity determines the drainage rate in the simplified 1D representation. In other words, the consideration of bedrock in the simplified 1D water table representation mimics the shallow water table obtained through the physics-based approach. In addition, the impact of the groundwater configuration is evaluated against the impact of the prescribed HFD using a small ensemble, which assesses the uncertainty of the subsurface initial condition. An analysis of variance indicates that both, groundwater configuration and HFD induce variability across the compartments of the terrestrial system, with decreasing impacts from the subsurface over the land surface towards the atmosphere. The induced variability, however, follows patterns of water table depths and HFDs.

The findings of this study confirm the hypothesis that groundwater dynamics contribute to land-atmosphere feedbacks at the continental scale. Moreover, results suggest that the consideration of groundwater dynamics in regional climate simulations might dampen the effect of land-atmosphere feedbacks in regions with a shallow water table. This effect may counteract the potential overestimation of land-atmosphere feedbacks in current regional climate models (Davin et al., 2016; Vautard et al., 2013). Furthermore, it is found that groundwater configurations and subsurface heterogeneity impact atmospheric variability, which emphasizes the importance of a realistic subsurface representation in atmospheric simulations.

Do atmospheric feedbacks to human water use affect water resources, and if so, to which extent might they contribute to continental drying (or wetting)?

The second part of this thesis evaluates the impact of human water use on local and non-local hydrologic and atmospheric states at the continental scale during the heatwave year 2003. Here, human water use is defined as groundwater abstraction and irrigation, and two-way feedbacks across the soil-vegetation-atmosphere system are simulated. In order to account for the uncertainty of human water use estimates and the associated land-atmosphere feedbacks, two data sets and two water use schedules are applied. The evaluation of the differences between the water use ensemble and a natural reference simulation enables the identification of time and space consistent feedback processes, which lead to a systematic change in the water cycle, considering input data uncertainty and internal model variability.

Simulation results show that atmospheric feedbacks to human water use may contribute to continental drying and wetting. Consistent spatio-temporal feedbacks and impacts across all water use scenarios, from monthly to annual and from grid-cell to regional scales, are found. Moreover, in this study, the simulated land-atmosphere feedbacks exceed net values

of human water use, i.e., irrigation minus groundwater pumping, at the watershed scale. Especially arid watersheds in Southern Europe suffer an additional loss of water through a decrease of the continental sink for atmospheric water (the difference between evapotranspiration and precipitation) in the simulations.

These results suggest that the impact of human water use transcends the local scale through changes of atmospheric water vapor transport, and indicate that effects of local human water use on water availability and sustainability may be of global importance via non-local feedbacks. These integrated feedbacks are not yet considered in regional climate simulations, reanalyses (Tuinenburg and de Vries, 2017), and water resource assessment studies (Jaramillo and Destouni, 2015), but constitute a potentially significant process and an additional uncertainty within the simulation of the terrestrial water cycle.

Does the incorporation of human water use in a continental-scale modeling system improve the simulation of evapotranspiration and precipitation?

The last part of this thesis addresses the added value of incorporating human water use in the simulation of the terrestrial cycle within a fully-coupled modeling system. The simulations from Chapter 4 are used to assess spatio-temporal model skills. Daily, *in-situ* and monthly, gridded observations of evapotranspiration and precipitation are used to evaluate the accuracy of these simulations at the local and watershed scale, respectively. The use of multiple gridded observational data sets enables an evaluation of human water use induced feedbacks against observational uncertainties at the watershed scale.

Results indicate that the incorporation of human water use significantly affects daily precipitation and evapotranspiration, and potentially improves the model skill through an improved simulation of local precipitation. A probabilistic analysis uncovers deficiencies in accurate and reliable simulations of precipitation, as evaluated by the occurrence of wet, heavy and very heavy precipitation days. Especially extreme precipitation events are overestimated, and the resulting wet biases are not removed through the incorporation of human water use. Moreover, the results do not conclusively show improved overall model skills. While the error at the daily scale is significantly reduced by the water use ensemble mean, watershed biases are found to be subject to large observational uncertainty. Using multiple gridded observational data sets, it is found that observations of evapotranspiration and precipitation exhibit an uncertainty that is larger than the land-atmosphere feedbacks induced by human water use, as simulated with TerrSysMP. Especially mountainous watersheds exhibit an increased wet bias, which is additionally associated with increased observational uncertainty.

The results of this study illustrate model deficiencies, which mainly arise from the simulation of precipitation. However, the identification of resulting biases and their propagation in fully-coupled modeling systems, remains challenging. In addition, model validation is challenged because observational uncertainty is high and the true errors remain unknown. This puts emphasis on the advantages of fully-coupled modeling systems, which do not rely on gridded atmospheric forcings for water assessment studies. These findings corroborate the need for extended observational data sets and further improvements of terrestrial water

cycle simulations.

Overall, the results from this thesis suggest that groundwater representations and human water use significantly affect water cycle processes beyond the catchment scale. The findings of this thesis highlight various aspects of the terrestrial water cycle, which remain underrepresented in current atmospheric and water assessment studies. In particular, groundwater dynamics might explain weak links in land surface and atmospheric models, which tend to overestimate land-atmosphere feedbacks, and impact atmospheric processes and stability beyond the boundary layer. Water assessment studies are typically driven with uncertain gridded atmospheric observations, and do not incorporate atmospheric feedbacks to changes in terrestrial hydrology. These limitations support the use of coupled modeling systems to advance the understanding of terrestrial water cycle processes and uncertainties. While the results illustrate the usefulness of continental-scale, fully-coupled groundwater to atmosphere simulations to study the combined, two-way feedbacks across the terrestrial system, several potential limitations require improvements of data sets and models to advance our understanding and the reliability of simulated water fluxes across the terrestrial system.

6.2   Limitations

In this thesis, the experimental model setup, the simulations and the findings are potentially limited by several factors. First, the high computational demand of fully-coupled simulations limits the findings of this thesis to the heatwave year 2003. The heatwave year 2003 was chosen, because the time period was characterized by an increased influence of the land surface and subsurface through water limitation and enhanced soil moisture-temperature feedbacks. Here, the performance of ensemble simulations (e.g., HFDs, initial conditions, human water use scenarios) was preferred over a single long-term simulation, in order to assess uncertainties in an expanded sensitivity framework. The findings in this thesis are thus representative of the extreme conditions during the heatwave in 2003.

Second, the simulations in this thesis were not calibrated with observations, and consequently exhibit a wet bias (as shown in Chapter 5). This wet bias is most likely connected to the high uncertainty of input parameters (e.g., hydraulic conductivity) and the initial conditions (e.g., water table depth), the relatively coarse resolution of all models, and the lack of the models to represent subgrid scale heterogeneity at these resolutions. Note in this context, that only overland flow is simulated. The coarse resolution of the surface and subsurface models potentially explains the wet bias in the subsurface (Shrestha et al., 2015). However, this is also connected to overestimated precipitation (cf. Chapter 5).

Moreover, the incorporation of human water use in the modeling system was static and limited to groundwater pumping (for industrial, agricultural and domestic use) and irrigation. Daily variations of groundwater pumping and irrigation were prescribed and did not account for water availability or water stress of plants. Water sources from, e.g., surface aquifers were neglected. Additionally, pumping wells were implemented at the grid cell level (i.e., one well on a 156 km^2 grid cell) and all reached a depth of 73 m (i.e., the center of the deepest soil layer).

Finally, the simulations did not account for feedbacks with other subsystems, such as the ocean, or human water use. Analogously, the seasonal cycle of vegetation was prescribed and carbon feedbacks were neglected. While these factors potentially confine the findings in this thesis to outcomes of sensitivity studies, validation with observations showed that the coupled model is able to realistically reproduce states and fluxes.

6.3 Future recommendations

This thesis contributed to the fundamental understanding of the terrestrial water cycle and related feedbacks from groundwater and human water use with land surface and atmospheric processes and opened the path for new research along technical and scientific aspects. For example, the setup of the coupled modeling system presented here should be regarded as a first step towards a process-based representation of the terrestrial water cycle in a continental-scale modeling system. Improvements, focusing on the existing modeling platform, could be made in three steps.

In a first step, the input data sets could be improved and updated, using, e.g., novel satellite based observations. With respect to the subsurface, the representation of a more detailed geology, e.g., with varying depths, is desirable. The uncertainty of the associated parameters and their feedbacks could further be addressed using an increased ensemble size. From the atmospheric perspective, the use of other convection parameterization schemes could be incorporated. The findings should furthermore be corroborated on longer, preferably decadal to centennial timescales. Finally, results should be evaluated against multiple observational data sets across the terrestrial system in order to identify model deficiencies.

In a second step, an increased resolution of all models is required. The resolution of the land surface and subsurface model could be increased in order to better resolve smaller river networks and smaller scale patterns of land cover and land use. Alternatively, a river channel parameterization might help to reduce the wet bias induced by the relatively coarse resolution. The effect of urban areas, which typically experience twice as much heat stress as their surroundings (Wouters et al., 2017), could be incorporated. Atmospheric simulations would benefit from an increased resolution, as this allows to resolve convection explicitly and avoids the use of erroneous parameterization schemes (e.g., Hohenegger et al., 2009; Hohenegger and Schär, 2007). Furthermore, the model could be calibrated with or constrained by observations. The application of data assimilation is advisable, and could help to deal with the high uncertainty of observational data sets. Multiple options emerge, including the assimilation of *in-situ* and remote sensing data, such as soil moisture (e.g., Kurtz et al., 2016; Naz et al., 2018), but also water storage anomalies from the Gravity Recovery and Climate Experiment, GRACE (e.g., Zaitchik et al., 2008), to improve the simulation of the terrestrial water cycle. Analogously, data assimilation could be used to support, e.g., subsurface parameter estimation where observations are scarce or missing (e.g., Baatz et al., 2017). At this stage, a dynamic formulation of human water use could be implemented (Wada et al., 2017; Condon and Maxwell, 2014b). Instead of prescribing water use and sources; one may apply a water allocation algorithm (Condon and Maxwell, 2013), which accounts for surface water and groundwater availability. A dynamic formulation

of human water use could potentially aggravate feedbacks especially during droughts and heatwaves.

Finally, the incorporation of other important feedbacks in the Earth system should be approached. The modular setup of the modeling platform allows to include additional processes and models. A few fundamental examples are provided in the following. An ocean model should be coupled to account for ocean-atmosphere feedbacks (Soden and Held, 2006), but also ocean-land feedbacks, such as sea-salt intrusion (Savenije, 1993), which significantly impacts water quality. Carbon cycle feedbacks, but also a better representation of vegetation and the water stress of plants could be achieved through the use of a dynamic vegetation model. Ultimately, interactions between human activities and terrestrial hydrology, e.g., through adjusted water use and reservoir regulation, but also pollution, should be considered (Wada et al., 2017).

The suggestions made here may be regarded as a proposed roadmap for future developments of an ESM based on, e.g., TerrSysMP. The development of a comprehensive ESM allows to study human and global impacts across climate subsystems, and should help to advance scientific knowledge on drivers and changes of the water cycle. Moreover, applications are not restricted to climate-related research. The use of coupled modeling systems could help to improve medium-range weather forecasts, but also predictions of water scarcity and flooding. However, all suggested approaches (adding model components, increasing resolutions and the application of data assimilation) increase the computational demand of this already extensive modeling approach. This challenges not only the computational power, but also storage and handling of the output of these models, and is hence emblematic of the big data issue (Hashem et al., 2015; Daniel, 2015). The overall aim, however, should be the creation of a terrestrial or Earth system reanalysis, which provides the best consistent estimate of the terrestrial system given observations and model physics. Ensemble simulations of past and future water fluxes across the climate system are required to adequately assess the uncertainty of water resources and develop adaptation strategies, which secure freshwater availability and socio-economic well-being.

Appendix

Appendix A

CAPE

The convective available potential energy (CAPE) is calculated according to Doswell and Rasmussen (1994):

$$\text{CAPE} = \int_{z_{LFC}}^{z_{EL}} g \left(\frac{T_{v,p} - T_{v,e}}{T_{v,e}} \right) dz \tag{A.1}$$

where g $[m/s^2]$ is the acceleration due to gravity, z_{LFC} $[m]$ is the height of the level of free convection, z_{EL} $[m]$ is the height of the equilibrium level, T_v $[K]$ is the virtual temperature of a specific air parcel p ascending with a constant pseudo-potential temperature in the environment e.

Appendix B

Analysis of Variance

The one-way analysis of variance (ANOVA, von Storch and Zwiers (2002)) is a generalization of the student t-test for more than two groups or treatments. The ANOVA is based on J samples of size n, represented by random variables Y_{ij} for $i = 1, \ldots, n$ and $j = 1, \ldots, J$. Here, j defines the sample and i the respective element of the sample. A statistical model

$$Y_{ij} = \mu + a_j + E_{ij} \tag{B.1}$$

is used to examine the variability of Y_{ij} due to J treatments. Here, $\mu = \frac{1}{J} \sum_{j=1}^{J} \mu_j$ denotes the overall mean for each sample mean μ_j. The coefficients $a_j = \mu_j - \mu$ describe the treatment effect and $E_{ij} \sim \mathcal{N}(0, \sigma^2)$ are the independent and identically distributed errors associated with the statistical model. The total variability of this model can then be decomposed into a treatment variability and the remaining variability. I.e. the total sum of squares

$$\mathrm{SST} = \sum_{i=1}^{n} \sum_{j=1}^{J} (Y_{ij} - Y_{oo})^2 \tag{B.2}$$

is decomposed into a treatment sum of squares

$$\mathrm{SSA} = n \sum_{j=1}^{J} (Y_{oj} - Y_{oo})^2 \tag{B.3}$$

and a sum of squared errors

$$\mathrm{SSE} = \sum_{j=1}^{J} \sum_{i=1}^{n} (Y_{ij} - Y_{oj})^2 \tag{B.4}$$

where the notation Y_{oo} denotes the unbiased estimator of μ and the subscript o indicates the average over the missing subscript. The ratio of SSA to SSE can be interpreted as a signal-to-noise ratio. Additionally, the portion of the total variance caused by the treatments, can be calculated and is called the adjusted coefficient of multiple determination

$$R^2 = \frac{\mathrm{SSA} - \frac{J-1}{J(n-1)} \mathrm{SSE}}{SST} \tag{B.5}$$

Large values of R^2 indicate that a large proportion of the response variable variance is induced by the treatments. Here, the ANOVA is performed for each grid cell and each

treatment separately. I.e., the variance caused by a specific treatment is tested against the total variance of the system (including all the other treatments). The statistical test of the estimated coefficients is based on the null hypothesis H_0 that there is no treatment effect and $H_0 : a_j = 0$ for all $j = 1, | \ldots, J$. If all coefficients $a_j = 0$ and consequently, H_0 is true, the ratio $f = \frac{(J(n-1))SSA}{(J-1)SSE}$ is distributed as a Fisher F random variable with $J-1, J(n-1)$ degrees of freedom specifying a significance level α. H_0 is accepted if $f < Q(1-\alpha, F(J-1, J(n-1)))$, where Q is the $(1 - \alpha)$ quantile of the Fisher F distribution.

Appendix C

Supplementary material to Chapter 3

This supporting information provides additional information of the impact of the experiment design. More specifically, it shows that the meso-scale circulation of the atmosphere is kept comparable between the single simulation experiments with a daily re-initialization of the atmosphere, as described in Sec. 3.3.4 in the manuscript. The supplementary material contains a descriptive text for three supplementary figures (Figures C.1- C.3), which show average pressure and wind speed differences between the simulations (TerrSysMP(3D)-TerrSysMP(FD)).

Daily re-initialization

In order to explain the rational for the daily re-initialization in more detail, differences of pressure and wind fields between TerrSysMP(3D) and TerrSysMP(FD) are analyzed. Figure C.1 shows the root mean square difference (RMSD) between atmospheric pressure in three model levels between TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1). In the highest level, around 550 hPa, the RMSD exhibits values of max. 6 Pa, but the differences are increasing towards the planetary boundary layer and close to the surface, hence indicating the increasing land surface impact. However, the RMSD is still relatively small with max. 30 Pa at the surface. In addition, the surface (and boundary-layer) differences follow some subsurface patterns, such as the Alpine region and channel of the Rhine river.

Moreover, juxtaposition of the time series of wind speed differences in Fig. C.2 and Fig. C.3 corroborates the findings. Significantly larger differences are modeled for TerrSysMP(HFD1) compared to TerrSysMP(HFD2), thereby indicating that circulation differences are partly induced by the different groundwater representations. The daily re-initialization at 00 UTC is indicated by a difference of 0 m/s. We conclude that the meso-scale structure of the atmosphere between the two simulations with different groundwater configurations is fairly similar and is not independent of the groundwater configuration. Hence, the impact of other sources, such as different circulations, contributing to the documented differences in the manuscript, is minimized and mostly local differences are analyzed.
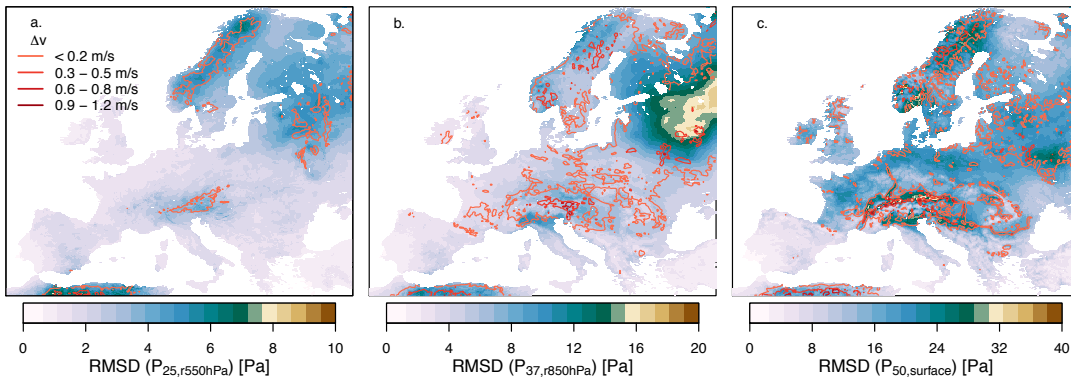
Figure C.1: Root mean square difference (RMSD) between the ensemble realization of TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1) over the entire simulation period based on hourly values of pressure (p [Pa]) for three heights: (a) model layer 25, reference pressure of 550 hPa, (b) model layer 37, reference pressure of 850 hPa and (c) lowest model layer representing surface pressure. Contour lines show the RMSD of wind speed between the ensemble of TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1).
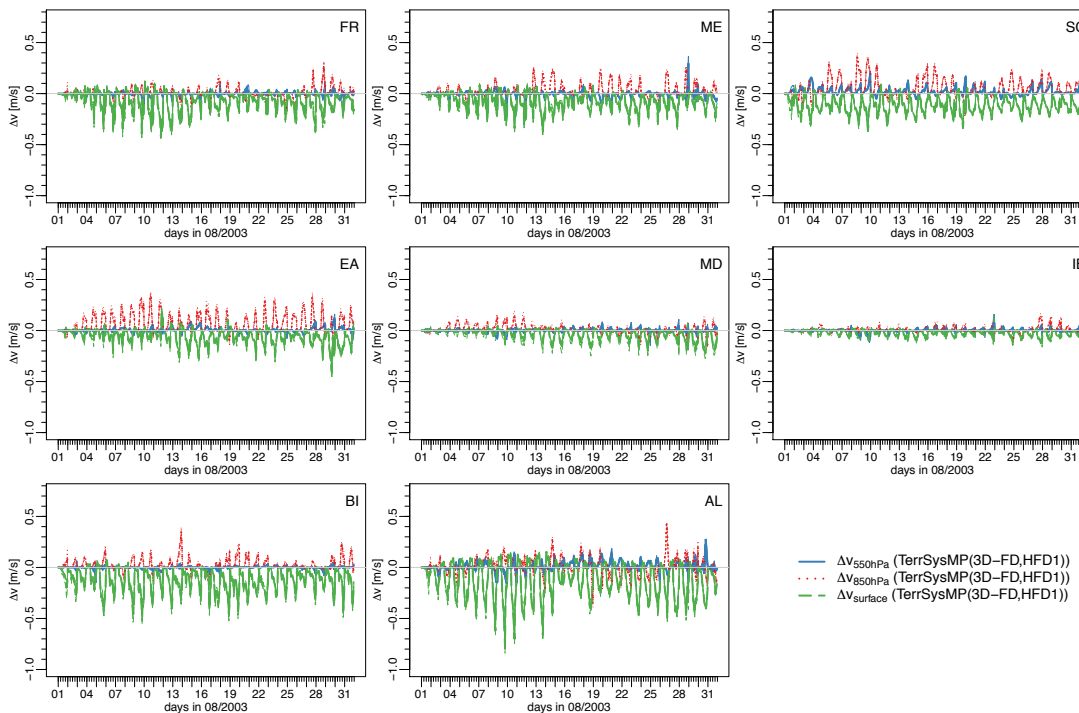
Figure C.2: Time series of wind speed differences between TerrSysMP(3D,HFD1) and TerrSysMP(FD,HFD1) for three heights (550 hPa, 850 hPa, surface), averaged over each PRUDENCE region (FR – France, ME – Mid-Europe, SC – Scandinavia, EA – Eastern Europe, MD – Mediterranean, IB – Iberian Peninsula, BI – British Islands, AL – Alpine region), respectively.
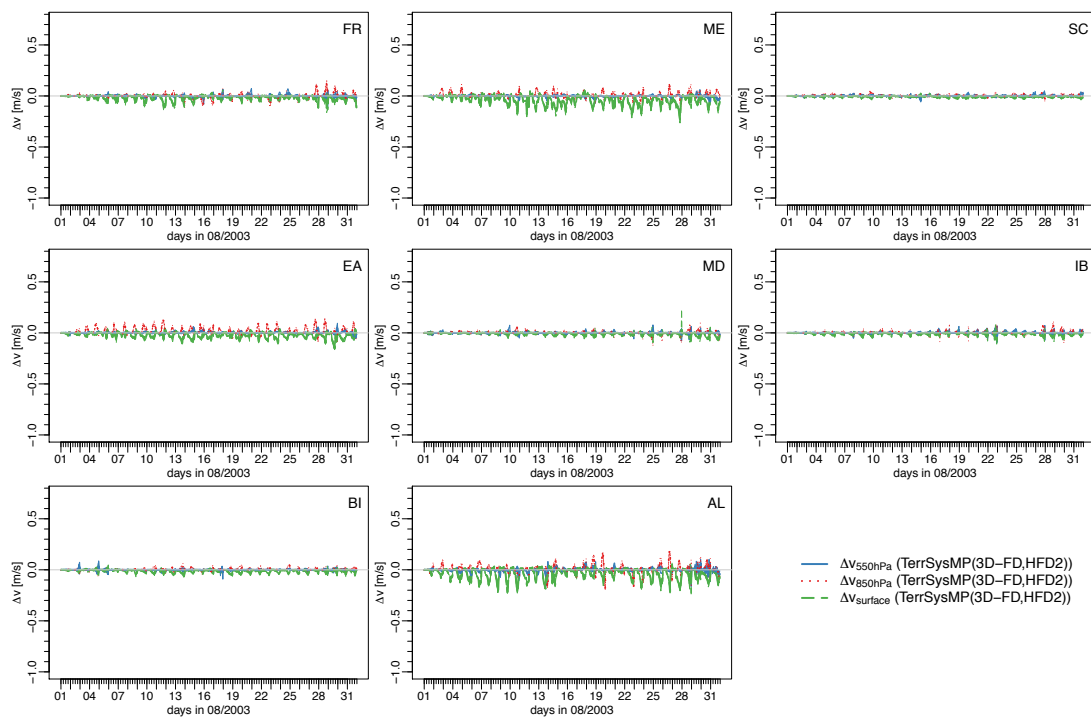
Figure C.3: Same as in Fig. C.2 but for HFD2.

Appendix D

Supplementary material to Chapter 4

This supporting information provides additional information for Chapter 4 on

1. The Modeling System

2. Land-atmosphere water balance

3. Computational costs

4. Human water use data

5. Results – supporting figures and tables

6. Validation of the simulations with *in-situ* observations.

Modeling System

In this study, we used the Terrestrial Systems Modeling Platform (TerrSysMP v1.2.0MCT, Shrestha et al., 2014; Gasper et al., 2014), which consists of the atmospheric model COSMO (Version 5.1, Baldauf et al., 2011; Doms and Schättler, 2002), the Community Land Model CLM (Version 3.5, Oleson et al., 2008) and the variably saturated surface-subsurface flow model ParFlow (Kollet and Maxwell, 2006; Jones and Woodward, 2001) coupled through OASIS-MCT (Valcke, 2013). TerrSysMP has already been set up over the European CORDEX domain at $0.11°$ ($\sim 12.5\ km$) resolution in previous studies (Keune et al., 2016) and simulates close-to-physics 3D groundwater dynamics and groundwater-surface water interactions, the land surface moisture, energy and momentum balance, and atmospheric transport in a fully-coupled fashion. In this study, COSMO5.1 was nested into 3-hourly boundary fields from ERA-Interim (Dee et al., 2011; $0.5°$ resolution). Additional 3, 6 and 9 h forecasts from ERA-Interim analysis at 00 and 12 UTC were used to update the lateral boundaries every 3 hours. The following 3D fields and geopotential are downscaled to $0.11°$ resolution and used for initialization and lateral boundaries: specific humidity, temperature, wind, cloud liquid, and ice water content. In order to keep the large-scale atmosphere consistent between all simulations, we applied the spectral nudging technique (von Storch et al., 2000) for horizontal wind components only (u,v) above the planetary boundary layer ($p < 850\ hPa$) for wave numbers smaller than 14 with $\alpha = 0.05$. We used the following physical parameterizations to account for subgrid-scale processes: Convection was parameterized using the Tiedtke mass flux scheme (Tiedtke, 1989). The δ-two-stream

approximation of the radiative transfer equation according to Ritter and Geleyn (1992) was used. Vertical turbulent diffusion was simulated with a 2.5-level closure scheme based on the prognostic equation for turbulent kinetic energy. Cloud micro-physics were simulated with a bulk- water continuity model, predicting cloud water, cloud ice, rain, and snow. COSMO's time step was 60 s, while CLM and ParFlow were applied with 180 s timesteps. Coupling took place every 180 s. The land surface characteristics for CLM3.5 were based on previous studies (Keune et al., 2016) and include land cover from the Moderate Resolution Imaging Spectroradiometer (MODIS, Friedl et al., 2002) data set, accounting for subgrid heterogeneity by employing a maximum of 4 plant functional types per grid cell. In addition, the leaf area index was updated using the MODIS 2000 data set. ParFlow was set up with the hydrofacies distribution 2 (HFD2) from Keune et al. (2016) in order to represent vertically heterogeneous soil and hydrogeologic characteristics. This HFD consists of the FAO database for the upper ten soil layers (reaching a depth of 3 m) and five bottom soil layers representing the deeper subsurface and bedrock hydrogeology using the Gleeson (Gleeson et al., 2011a,b) database. The vertical discretization increases with depth to reach a total of 103 m subsurface depths. At the coastlines, a lateral Dirichlet boundary condition was applied with a hydrostatic profile, which mimics the ocean with a free water table at the top. Five simulations were carried out covering the full calendar year 2003. In addition to four human water use simulations, one natural reference simulation (NAT) was performed, which does not consider any human water use. The water use simulations were constructed as follows in order to account for the uncertainty of daily pumping and irrigation estimates applied in the simulations, and the atmospheric feedback by perturbing the water use schedule. Irrigation and (total) groundwater abstraction estimates were taken from Wada et al. (2012), Wada et al. (2016a) and Siebert et al. (2010), Siebert and Döll (2010). Total groundwater abstraction (abstraction for irrigation, domestic and industrial demand) was applied, and estimates of domestic and industrial demand from Wada et al. (2012), Wada et al. (2016a) were added to groundwater abstraction estimates for irrigation from Siebert et al. (2010) and Siebert and Döll (2010) for consistency. Note that while we did not represent surface water abstraction explicitly, total irrigation includes water from surface water and groundwater in all model simulations. The two data sets with an original resolution of 0.1° were bilinearly interpolated in space to match the CORDEX grid. Two water use schedules, i.e., daytime (06:00–18:00 UTC) and nighttime (18:00 - 06:00 UTC) were generated for each data set, resulting in a total of four water use scenarios. The hourly rates of each daily estimate were added to the top soil layer (irrigation) or subtracted from the bottom soil layer (groundwater abstraction/pumping) in ParFlow via straight-forward source/sink terms in the governing equations. Irrigation and pumping were applied simultaneously in the model. All simulations started from the same quasi-equilibrium initial condition, which was achieved by a multi-year spin-up of the NAT simulation using atmospheric forcing from the year 2003. The first 5 days of each simulation were neglected for analysis.

Land-atmosphere water balance

In a natural system, the terrestrial hydrologic cycle extends from the subsurface over the land surface to the atmosphere, connecting the land surface/subsurface water balance

$$(\partial S/\partial t)_{NAT} = P_{NAT} - ET_{NAT} - (div(Q_g))_{NAT} - R_{NAT} \tag{D.1}$$

with the atmospheric water balance

$$(\partial W/\partial t)_{NAT} = -P_{NAT} + ET_{NAT} - (div(Q))_{NAT} \tag{D.2}$$

via precipitation P and evapotranspiration ET. Terrestrial water storage changes arise through the balance of P, ET, runoff (R) and groundwater divergence $div(Q_g) = \nabla \cdot Q_g$. Vice versa, the atmospheric moisture storage can only be altered through P, ET and the atmospheric moisture divergence $div(Q)$. The atmospheric moisture divergence $div(Q)$ is the vertically integrated horizontal moisture transport defined as

$$div(Q) = -\frac{1}{g}\nabla \cdot \int_0^{p_s} qv_h dp, \tag{D.3}$$

with the gravitational acceleration g, specific humidity q, horizontal wind vector v_h, pressure p and surface pressure p_s. As the atmospheric water storage is relatively small and its temporal change is even smaller ($(\partial W/\partial t) \approx 0$) even over relatively short time scales, the atmospheric moisture divergence $div(Q)$ is mainly driven by P and ET, i.e.,

$$div(Q) = ET - P. \tag{D.4}$$

We define continental sink C_{SI} and continental source C_{SO} based on the atmospheric divergence, i.e., $C_{SI} = -div(Q) = P - ET$ for $P > ET$ and $C_{SO} = div(Q) = ET - P$ for $P < ET$. Furthermore, substitution of Eq. (D.1) into Eq. (D.4) combines the atmospheric and land surface/subsurface water balances in a natural system

$$(div(Q))_{NAT} = -(\partial S/\partial t)_{NAT} - (div(Q_g))_{NAT} - R_{NAT} \tag{D.5}$$

and shows that the land surface/subsurface act as a sink/source for atmospheric water. We calculate $div(Q)$ as the residual of Eq. (D.2) based on hourly outputs of P and $(\partial W/\partial t)$ from COSMO, and ET from CLM, respectively. Figure 4.1 of the main manuscript illustrates simulated cumulative $(div(Q))_{NAT}$, for each grid cell over a year, the summer months (June-July-August), and for August. For the watershed calculation of $div(Q)$, we used the watershed boundaries from HydroSHEDS (Global Hydrological data and maps based on Shuttle Derivatives at multiple Scales, Lehner et al., 2006) at 30 arc seconds ($\sim 1\ km$). $div(Q)$ was summarized over the year, summer months and August and the average of all grid cell values within each watershed was calculated. The natural hydrologic cycle, as described above, is systematically altered through human water use (HWU), here considered as groundwater pumping G and irrigation I. Hence, in a managed system, the natural land

surface balance is disturbed

$$(\partial S/\partial t)_{HWU} = P_{HWU} - ET_{HWU} - (div(Q_g))_{HWU} - R_{HWU} \qquad (D.6)$$

and might exhibit systematic differences compared to the natural reference system in evapotranspiration

$$\Delta ET = ET_{HWU} - ET_{NAT}, \qquad (D.7)$$

in precipitation,

$$\Delta P = P_{HWU} - P_{NAT}, \qquad (D.8)$$

and, thus, in atmospheric divergence,

$$\Delta(div(Q)) = (div(Q))_{HWU} - (div(Q))_{NAT}, \qquad (D.9)$$

and subsurface water storage,

$$\Delta(\partial S/\partial t) = ((\partial S/\partial t))_{HWU} - ((\partial S/\partial t))_{NAT}. \qquad (D.10)$$

The deviations illustrated here account for the integrated, non-linear feedbacks of the system, i.e., ΔET accounts for the difference in ET induced by $I - G$, but also for the ET feedbacks on P, which in turn affect ET in a non-linear feedback loop. $div(Q)$ was calculated for each human water use scenario separately and averaged over all grid cells within each watershed. The difference between the natural and human water use were calculated as described in Eq. (D.7) to Eq. (D.10). For the consistency check, $div(Q)$ values were averaged over all terrestrial grid cells in each PRUDENCE (Christensen et al., 2007) region. The PRUDENCE regions have the following extensions (W,E,S,N): Iberian Peninsula, IB (-10,3,36,44); British Isles, BI (- 10,2,50,59); Mid-Europe, ME (2,16,48,55); Mediterranean, MD (3,25,26,44); Eastern Europe, EA (16,30,44,55), Alps, AL (5,15,44,48), France, FR (-5,5,44,50); and Scandinavia, SC (5,30,55,70). For the CORDEX domain average, a sponge zone was removed from the domain and the average over all terrestrial grid cells in a focus domain (Fig. 4.1) was calculated. To evaluate the consistency, the total annual, summer and August differences $\Delta(div(Q))_{HWU1}$ were plotted against $\Delta(div(Q))_{HWU2}$, averaged over the PRUDENCE regions, the watersheds and single grid points. As described in Eq. (D.6) and Eq. (D.10), the human water use induced feedbacks contribute to drying or wetting. We calculate the subsurface storage S for each simulation, over all terrestrial grid cells and the full soil column of each cell, containing the fully saturated and the variably saturated zones. For one cell, the storage is the sum of water in each layer k of thickness Δz_k, i.e.,

$$S = \sum_{k=1}^{K} [s_k \phi \Delta z_k], \qquad (D.11)$$

with relative saturation s, and porosity ϕ. Because all simulations start with the same initial terrestrial water storage, we can directly compare the total storages of the natural

and managed systems at the end of the simulations, i.e.,

$$\Delta S = \Delta(\partial S/\partial t) = (\partial S/\partial t)_{HWU} - (\partial S/\partial t)_{NAT} = S_{end,HWU} - S_{end,NAT} \qquad \text{(D.12)}$$

and analyze how human water use induced changes of the continental sink $\Delta(div(Q))$ translate into water storage changes according to

$$\Delta(\partial S/\partial t) = \Delta(div(Q)) - [\Delta(div(Q_g)) + \Delta R] \qquad \text{(D.13)}$$

A storage change according to Eq. (D.12) was calculated for each grid cell and averaged over the grid cells in each watershed. To illustrate the relation of ΔS and $\Delta(div(Q))$ from Eq. (D.13), we calculated annual $\Delta(div(Q))$ averaged over each watershed and plotted the values against ΔS. Deviations from the 1:1 line in Fig. 4.4 of the main manuscript are thus due to changes in lateral surface-subsurface flow, i.e., groundwater divergence and runoff $[\Delta(div(Q_g)) + \Delta R]$ in Eq. (D.13). For visualization, no changes ($\Delta S = 0$ or $\Delta(div(Q)) = 0$) were neglected in Fig. 4.4 and the slope of the relationship changed because of sign changes for logarithmic scaling.
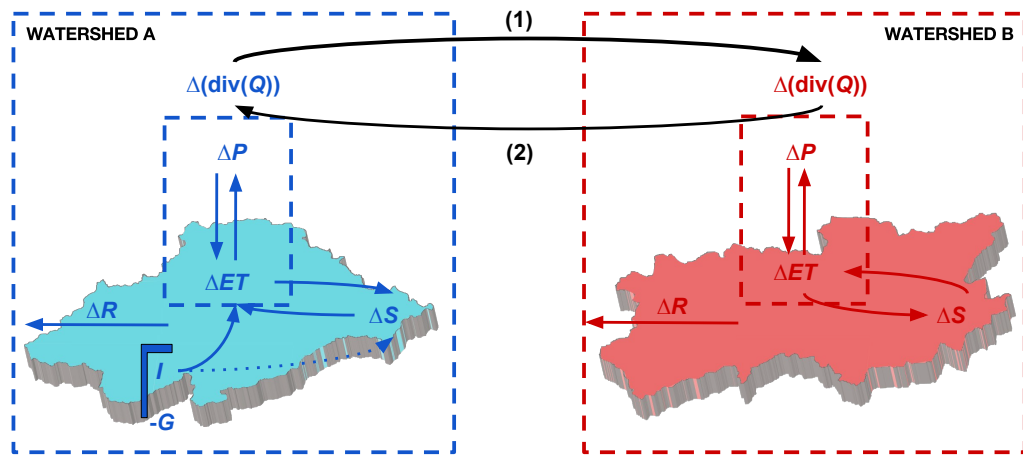
Figure D.1: How human water use can alter the continental sink of moisture across watersheds. In a managed watershed A the natural water balance is disturbed through human water use (groundwater pumping G and irrigation I) induced changes of water storage ΔS, runoff ΔR, evapotranspiration ΔET and precipitation ΔP. First, local feedbacks to human water use alter the continental sink of watershed A, and the atmospheric water vapor transport $div(Q)$ between remote watersheds A and B, which can trigger or inhibit precipitation initiation in watershed B. Secondly, these remotely induced changes of the continental sink and water storage of watershed B can further feedback to watershed A (two-way feedback).

Computational costs

All simulations were carried out on the Cray XC40 European Centre for Medium Range Weather Forecasts (ECMWF). Each simulation was run on 14 nodes (with each 36 tasks, i.e., 504 total tasks), split into 12 x 12 (144) tasks for ParFlow, 6 x 6 (36) tasks for CLM3.5 and 18 x 18 (324) tasks for COSMO. One year of simulation required on high-performance computing system at the average about 1.223.067 SBUs ($\sim$75.919 core-hours).

Human water use data
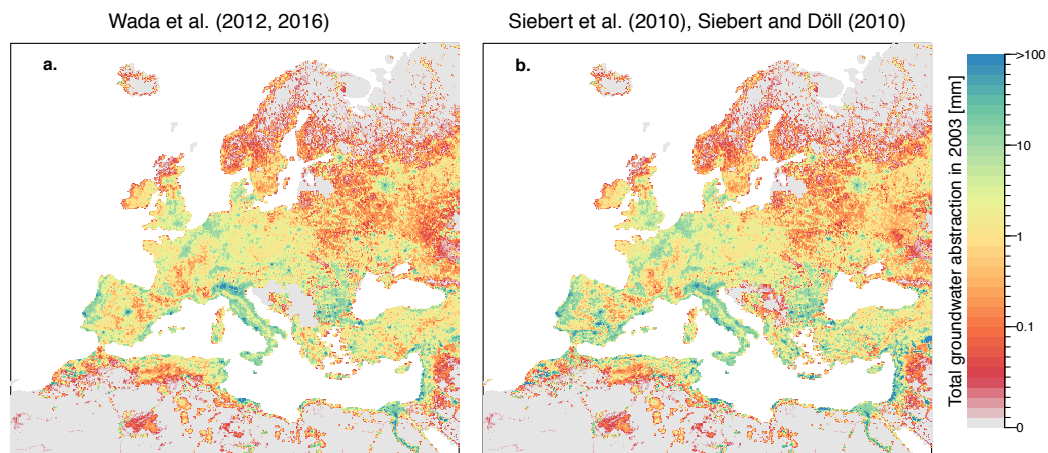
Figure D.2: Total groundwater abstraction [mm] from both data sets, (a) HWU1, Wada et al. (2012, 2016a) and (b) HWU2, Siebert et al. (2010), Siebert and Döll (2010), interpolated to the European CORDEX domain at 0.11° resolution. Note that groundwater abstraction for industrial and domestic use from Wada et al. (2012, 2016a) were added to groundwater abstraction from Siebert et al. (2010), Siebert and Döll (2010) for consistency.
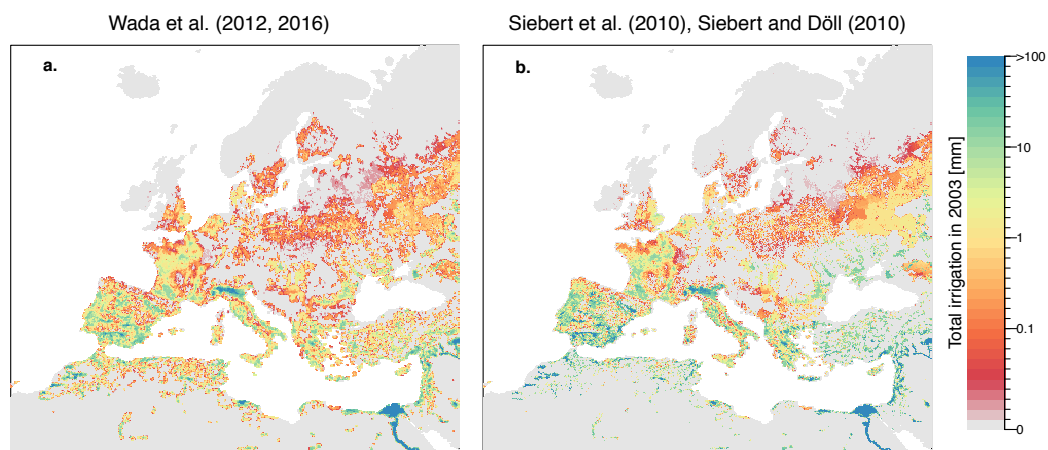
Figure D.3: Total irrigation [mm] from both data sets, (a) HWU1, Wada et al. (2012, 2016a) and (b) HWU2, Siebert et al. (2010), Siebert and Döll (2010), bilinearly interpolated to the European CORDEX domain at 0.11° resolution.
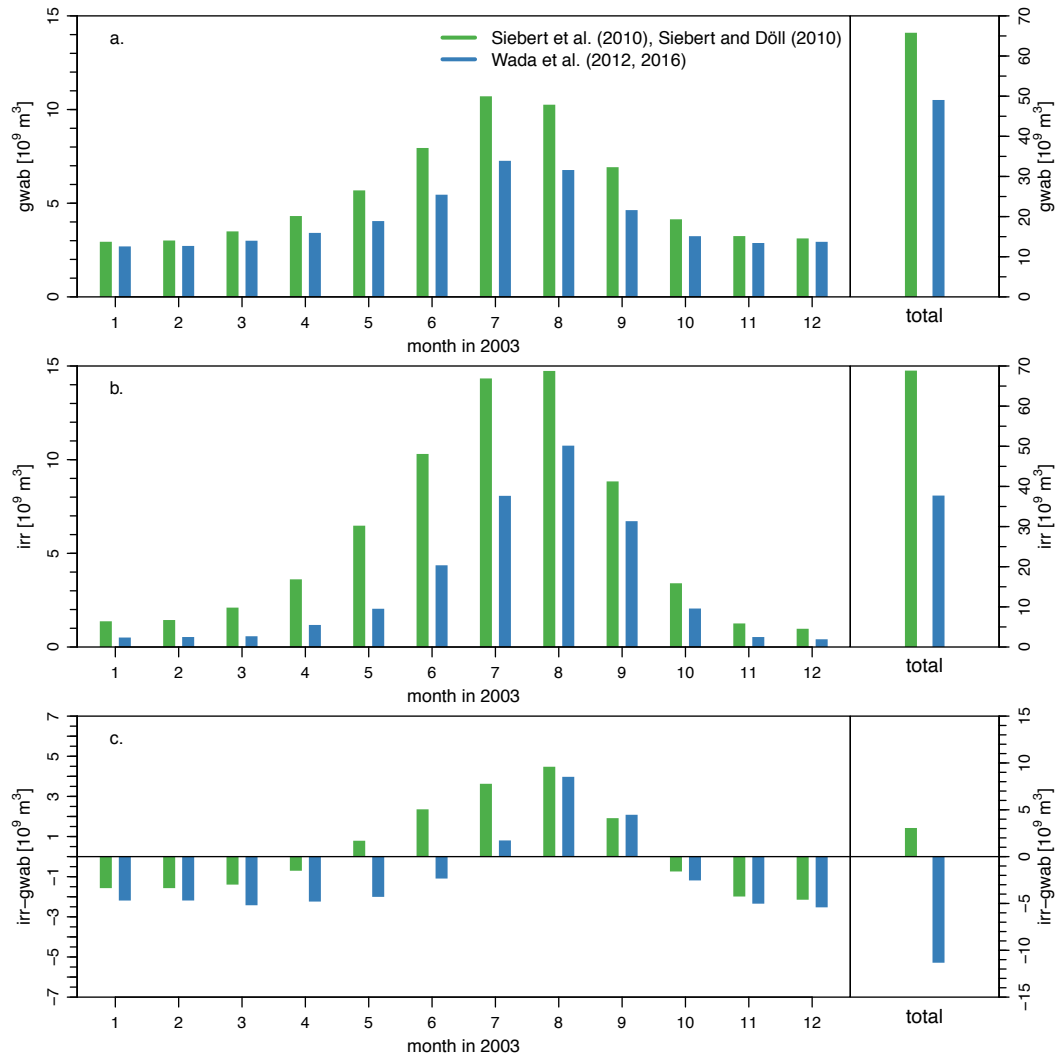
Figure D.4: Monthly sums of (a) groundwater abstraction [m3], (b) irrigation [m3] and (c) their differences as (irrigation-groundwater abstraction) for both data sets (blue: HWU1, i.e., Wada et al. (2012, 2016a), green: HWU2, i.e., Siebert et al. (2010), Siebert and Döll (2010)) over the European CORDEX domain for the year 2003.

Results - supporting figures and tables

Table D.1: Atmospheric divergence ($div(Q)$) simulated with the natural reference run (NAT) and the 4 water use scenarios (HWU1-1, HWU1-2, HWU2-1, HWU2-2) over the entire year 2003, averaged over the land surface of the entire CORDEX domain. The last column indicates whether the difference between HWU and NAT is significant at the 95% confidence interval (CI; * = significantly different) using all grid points in a paired t-test (df=50303).

	div(Q) [mm]	$\Delta(div(Q))_{HWU-NAT}$ [mm] (95% CI)	$\Delta(div(Q))_{HWU-NAT}$ [%]	Significant difference (t-value, p-value)
NAT	-459.88	-	-	-
HWU1-1	-459.09	+0.78 (0.24-1.33)	-0.17	* (t=2.83, p<0.05)
HWU1-2	-456.64	+3.23 (2.70-3.76)	-0.70	* (t=11.91, p<0.001)
HWU2-1	-458.27	+1.60 (1.07-2.14)	-0.35	* (t=5.91, p<0.001)
HWU2-2	-457.80	+2.08 (1.55-2.61)	-0.45	* (t=7.67, p<0.001)
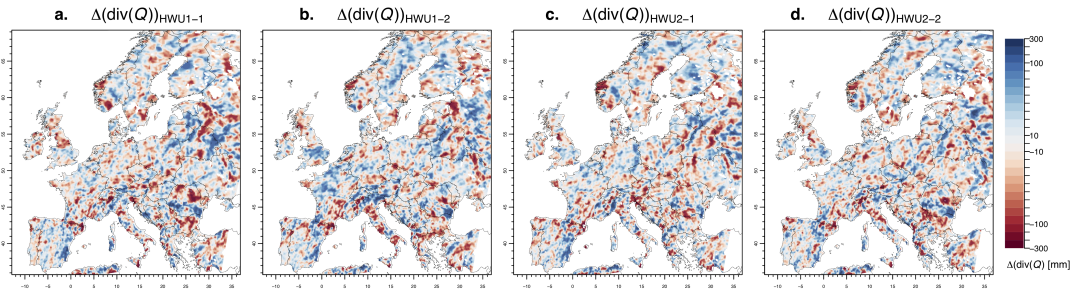
Figure D.5: $\Delta(div(Q))_{HWU-NAT}$ [mm] over the full CORDEX domain for all water use scenarios (a. HWU1-1, b. HWU1-2, c. HWU2-1, d. HWU2-2) over the full year 2003.
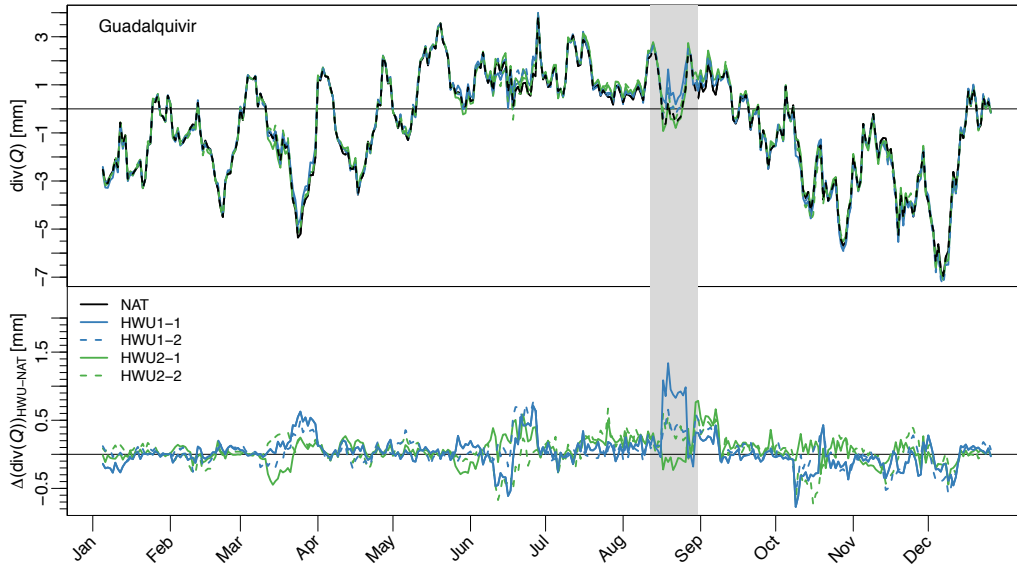
Figure D.6: Time series of $div(Q)$ [mm] and $\Delta(div(Q))$ [mm] averaged over the Guadalquivir basin (Iberian Peninsula). A 10-day-running mean was used for better illustration. Negative values of $div(Q)$ illustrate that the land is a sink of moisture, and positive values illustrate that the land provides moisture to the atmosphere. Note that human water use can potentially reverse this balance as illustrated in August by HWU1-1.
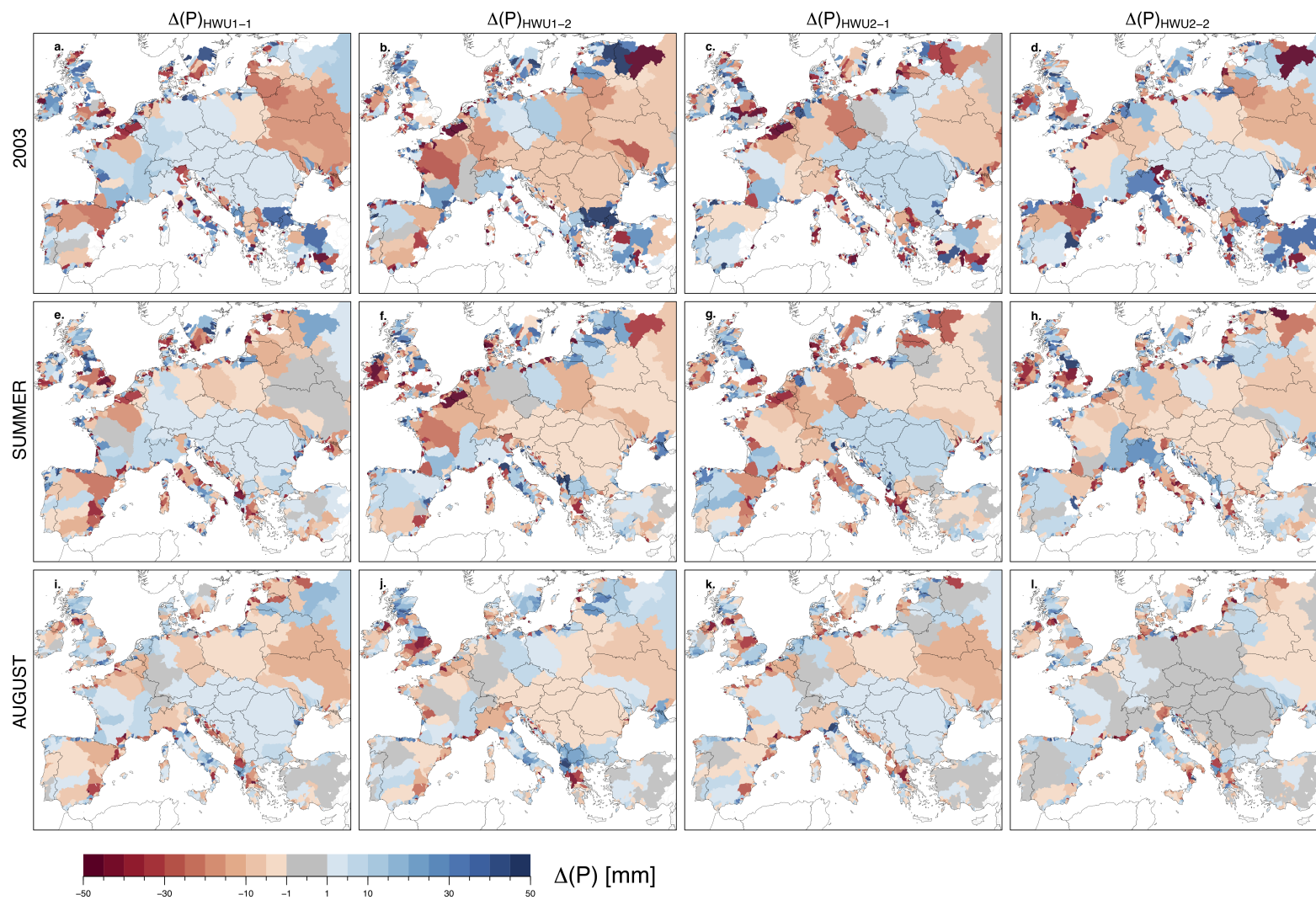
Figure D.7: Human water use-induced amplification or attenuation of precipitation (ΔP [mm]) over European watersheds for all water use scenarios along the columns (HWU1-1 (a,e,i), HWU1-2 (b,f,j), HWU2-1 (c,g,k), HWU2-2 (d,h,l)) over the full year 2003 (a-d), summer 2003 (JJA, e-h), and August 2003 (i-l). Watersheds with P differences between -1 mm and 1 mm are marked grey.
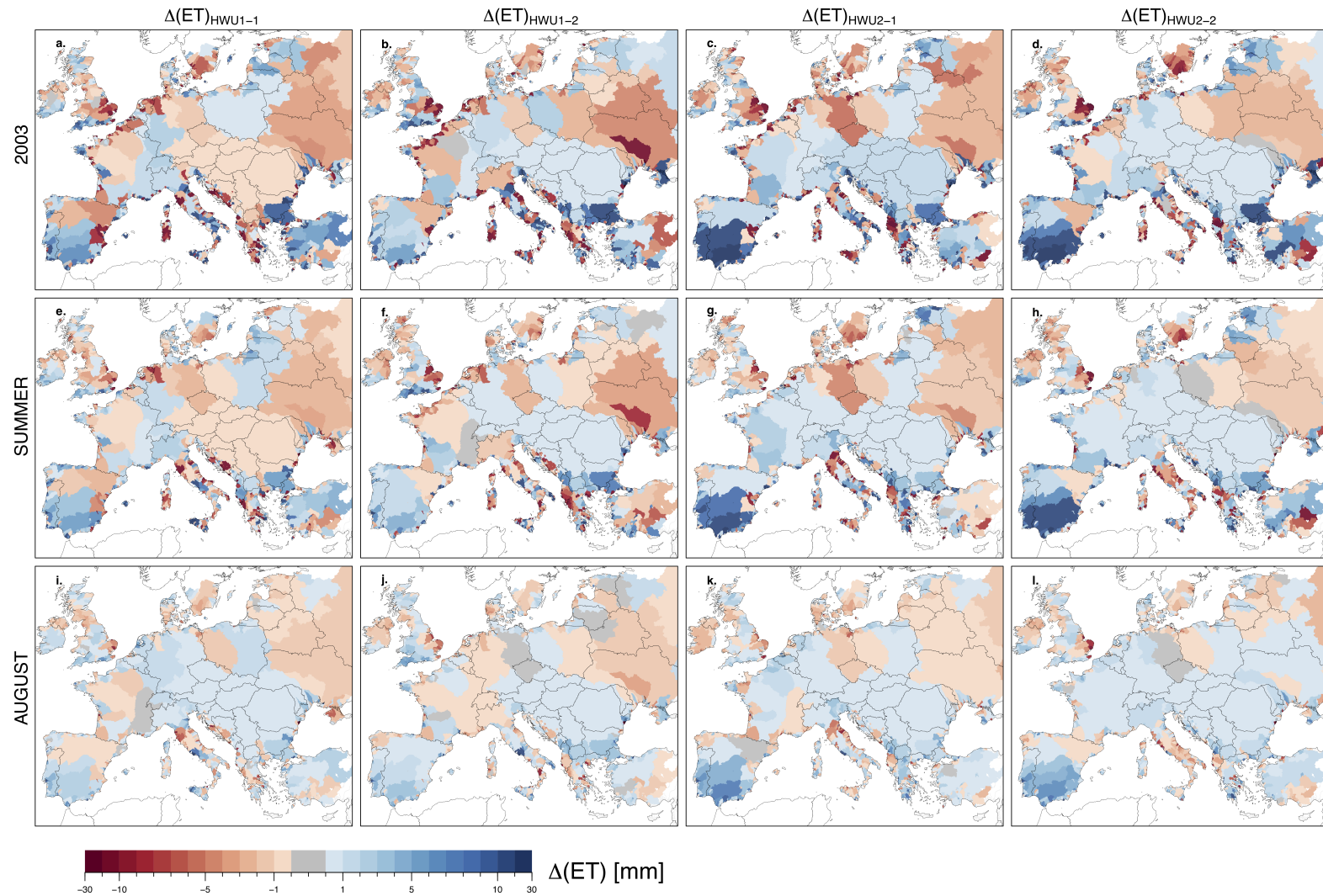
Figure D.8: Simulated human water use-induced amplification or attenuation of evapotranspiration (ΔET [mm] as HWU-NAT) over European watersheds for all water use scenarios along the columns (HWU1-1 (a,e,i), HWU1-2 (b,f,j), HWU2-1 (c,g,k), HWU2-2 (d,h,l)) over the full year 2003 (a-d), summer 2003 (JJA, e-h), and August 2003 (i-l). Watersheds with ET differences between -0.1 mm and 0.1 mm are marked grey.
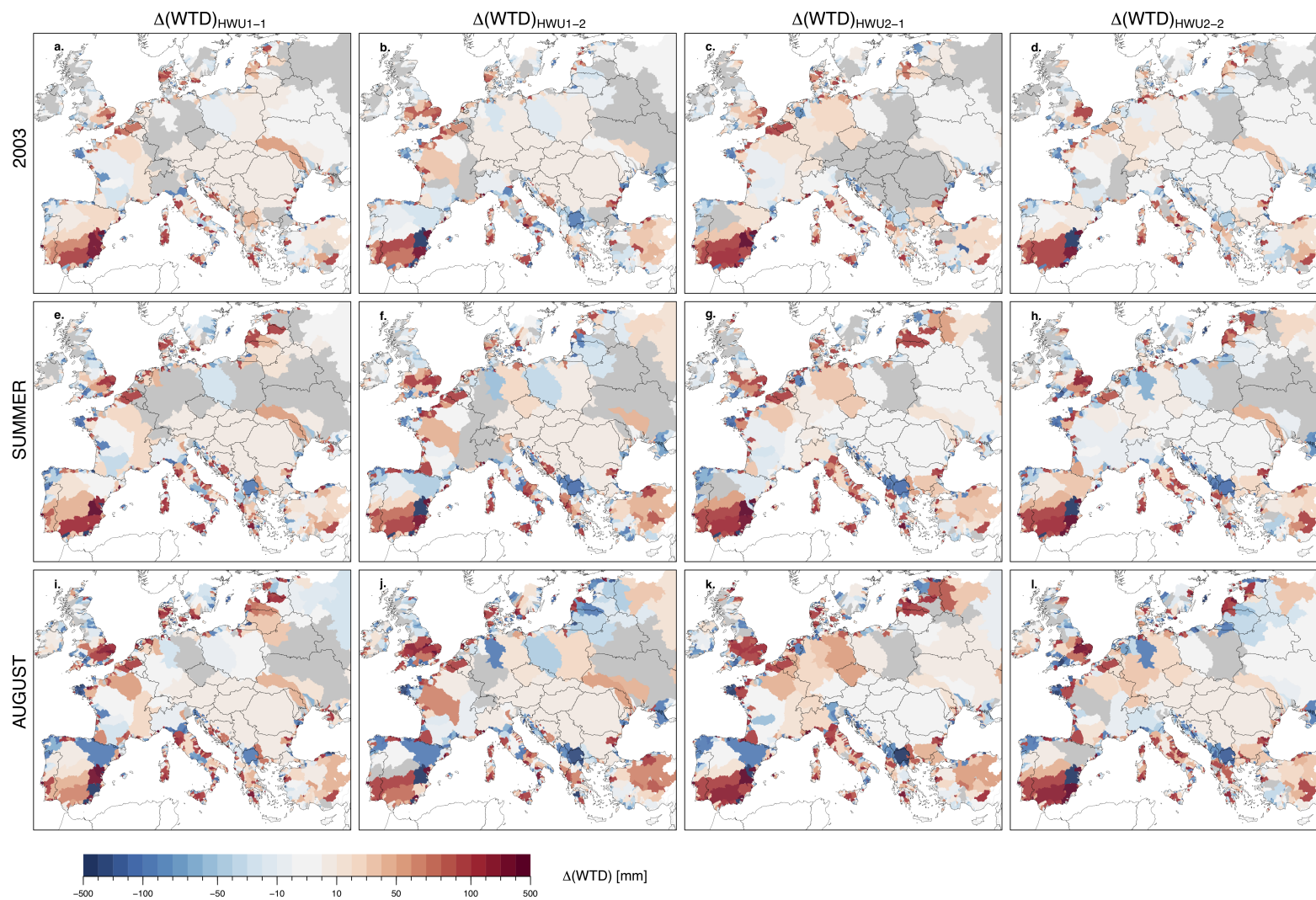
Figure D.9: Simulated human water use-induced groundwater table rise (blue) or decline (red) (ΔWTD [mm]) over European watersheds for all water use scenarios along the columns (HWU1-1 (a,e,i), HWU1-2 (b,f,j), HWU2-1 (c,g,k), HWU2-2 (d,h,l)) over the full year 2003 (a-d), summer 2003 (JJA, e-h), and August 2003 (i-l). WTD differences between -5 mm and 5 mm are masked grey.
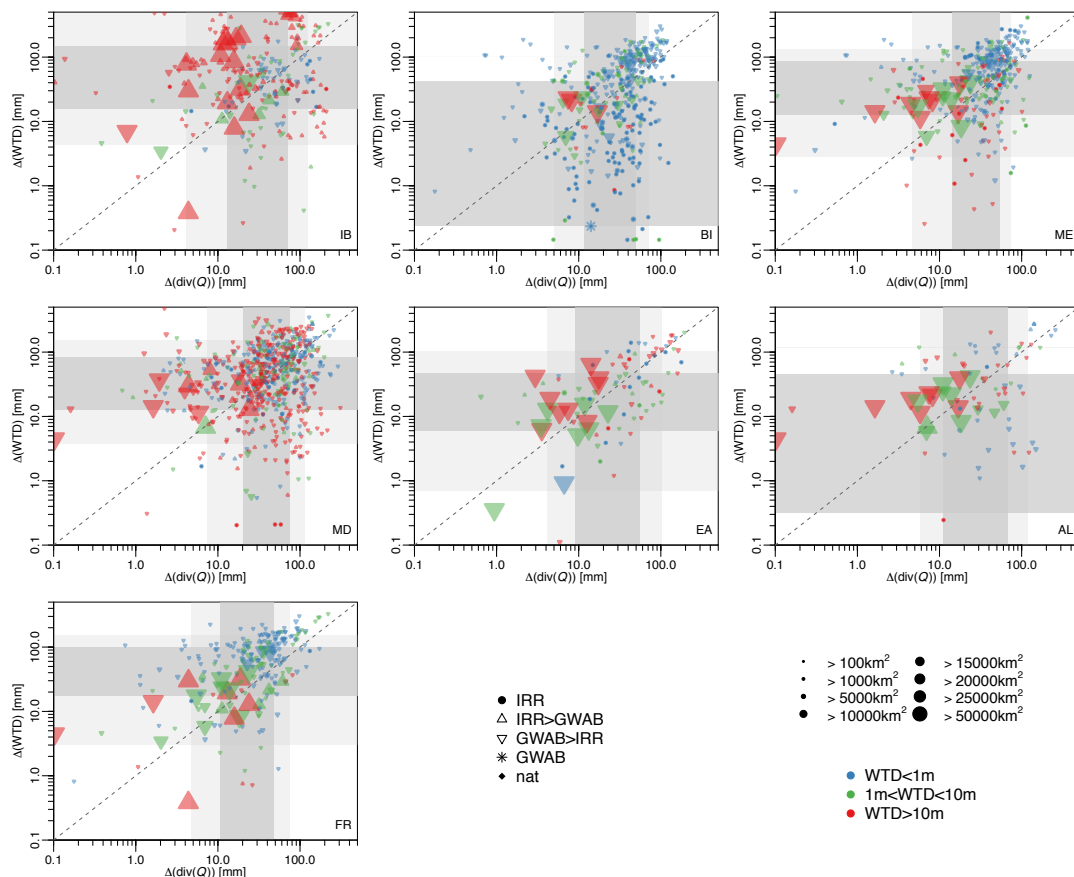
Figure D.10: Water table declines ΔWTD [mm] as a function of changes of the atmospheric sink $\Delta(div(Q))$ [mm] for all watersheds in the PRUDENCE regions IP (Iberian Peninsula), BI (British Isles), ME (Mid-Europe), MD (Mediterranean), EA (Eastern Europe), AL (Alps) and FR (France). The colors represent the reference water table depth (blue: shallow $(0-1\ m)$, green: medium $(1-5\ m)$, red: deep $(>5\ m)$). The size of the symbols is commensurate with the watershed size. The symbols illustrate the most dominant water use. All water use scenarios considered.

Validation of the simulations with *in-situ* observations

We validated daily precipitation and daily evapotranspiration from all simulations (NAT, HWU1-1, HWU1-2, HWU2-1, HWU2-2) with rain gauge observations from the European Climate Assessment & Data (ECA&D, http://www.ecad.eu) data set, and with eddy covariance measurements from FLUXNET (Baldocchi et al., 2001, http://fluxnet.fluxdata.org). The half- hourly values of FLUXNET are accumulated to daily latent heat fluxes and converted to daily evapotranspiration (excluding missing values consistently from model results). For the validation with rain gauge observations, simulations are interpolated to the stations using the nearest neighbor method. A total of 1246 stations are used for validation[1]. For the comparison with FLUXNET stations, the nearest neighbor is used irrespective of the prevalent land cover. This allows us to assess the accuracy of evapotranspiration at 28 towers across Europe with varying land cover types i.e., evergreen broadleaf forest (EBF, 4 towers); deciduous broadleaf forest (DBF, 7 towers); evergreen needleleaf forest (ENF, 6 towers); mixed forest (MF, 1 tower); grassland (GRA, 9 towers) and crop (CRO, 1 station).
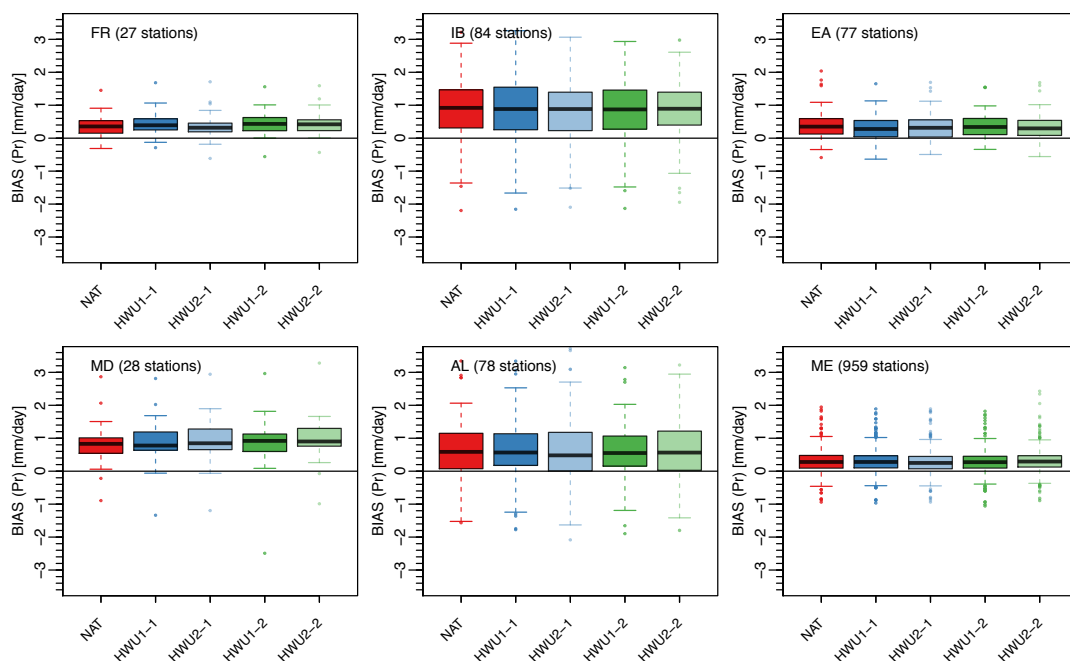
Figure D.11: Box-whisker plots of the daily precipitation bias (simulations - observation) in *mm/day* for all stations within the respective PRUDENCE regions, evaluated with ECA&D stations over the entire year 2003. The number in the top of each subplot shows the total number of stations available for that period.

Figure D.11, as shown above, shows the annual precipitation bias [*mm/day*] for all ECA&D rain gauge observations over the PRUDENCE regions France (FR), Iberian Peninsula (IP), Eastern Europe (EA), Mediterranean (MD), Alps (AL) and Mid-Europe (ME). The figure shows that all simulations exhibit a wet bias from 0.2 to 1 *mm/day*. Especially southern Europe (the Mediterranean and the Iberian Peninsula), as well as the Alps show a wet bias on the order of 1 *mm/day*. This agrees well with common regional climate model validations (e.g., Kotlarksi et al., 2017; Katragkou et al., 2015). However, precipitation in Mid- and Eastern Europe, as well as France is more accurate and shows only a small, but wet bias. Vice versa, summer precipitation can be underestimated, as shown in Fig. D.12. Summer precipitation is underestimated for all simulations over the Iberian Peninsula and the Mediterranean (−1 *mm/day*) and has only a small wet bias for the rest

---

[1]Note that all observations are used, irrespective of quality flags and data availability

of Europe. The difference in the mean biases between all simulations is comparably small and does not allow for any skill assessment.
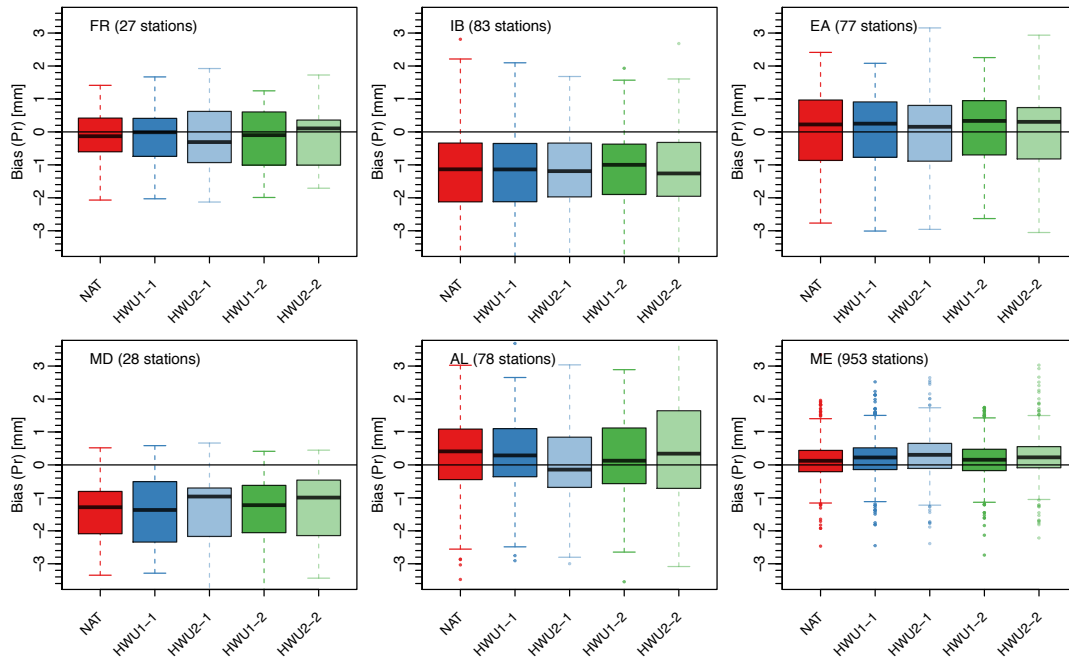
Figure D.12: Box-whisker plots of the daily summer precipitation bias (simulations - observation) in mm/day for all stations within the respective PRUDENCE regions, evaluated with ECA&D stations over June-July-August 2003. The number in the top of each subplot shows the total number of stations available for that period.

This tendency for overestimated precipitation may in turn lead to an overestimation of evapotranspiration. The location of all FLUXNET towers used for validation is shown in Fig. D.13 and Fig. D.14 shows the annual and summer evapotranspiration biases at these stations. However, there is no clear tendency visible at the available towers. Nevertheless, some stations, such as FR- Puc, IT-Cpz, IT-on, IT-SRo and PT-Mi1 are clearly overestimating evapotranspiration. Interestingly, these are mainly southern (IT, PT) or heatwave affected locations (FR), which might indicate that the simulations are not water-limited enough (potentially arising from the wet precipitation bias). A similar conclusion can be drawn from the summer evaporation bias. Nevertheless, there are also a few stations which underestimate evapotranspiration and/or have only a comparably small bias ($< \pm1\ mm/day$). However, there is no clear pattern with respect to the land use type visible. Considering also the observational uncertainty, the large areal representativeness of the simulated evapotranspiration in contrast to the small footprint of the tower, and the fact that we did not filter for the land use type, indicates that the simulations agree reasonably well with the observations with a tendency for overestimating evapotranspiration. This has also been shown in previous studies and strongly depends on the spatial resolution of the coupled modeling system (Shrestha et al., 2015). Again, all simulations perform similarly, but there are some differences for e.g. summer evapotranspiration at IT-Non, HU-Bug and IT-Ro1 and IT-Ro2 visible, which indicate an impact of human water management.
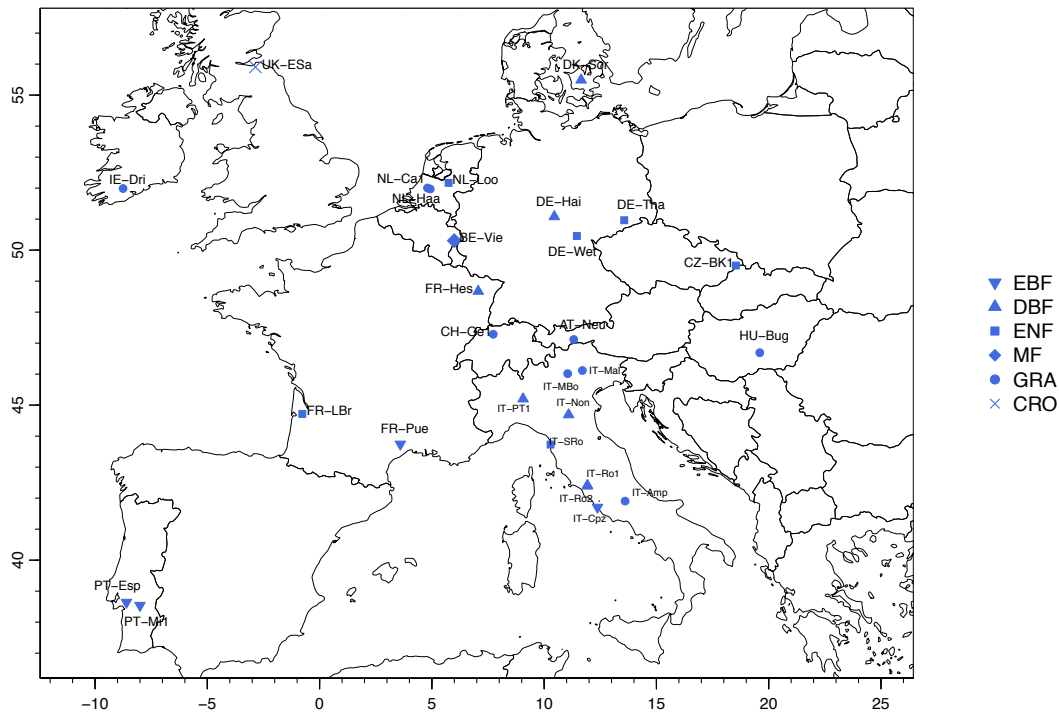
Figure D.13: Locations of FLUXNET towers available in 2003 (*ET*). The symbols indicate the dominant land cover.
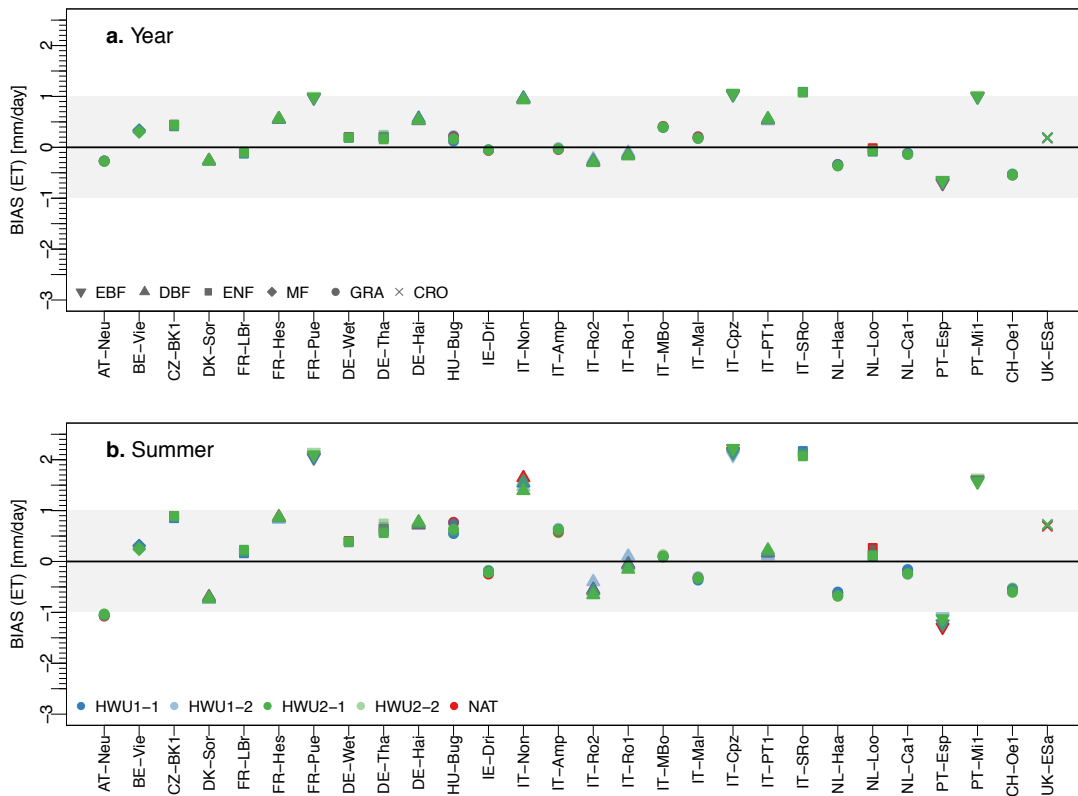
Figure D.14: Daily evapotranspiration bias (simulations - observation) in *mm/day* for all available FLUXNET stations for (a) the entire year 2003 and (b) June-July-August 2003. Colors represent the simulation scenario and symbols indicate the dominant land use type of the FLUXNET tower.

Appendix E

Supplementary material to Chapter 5
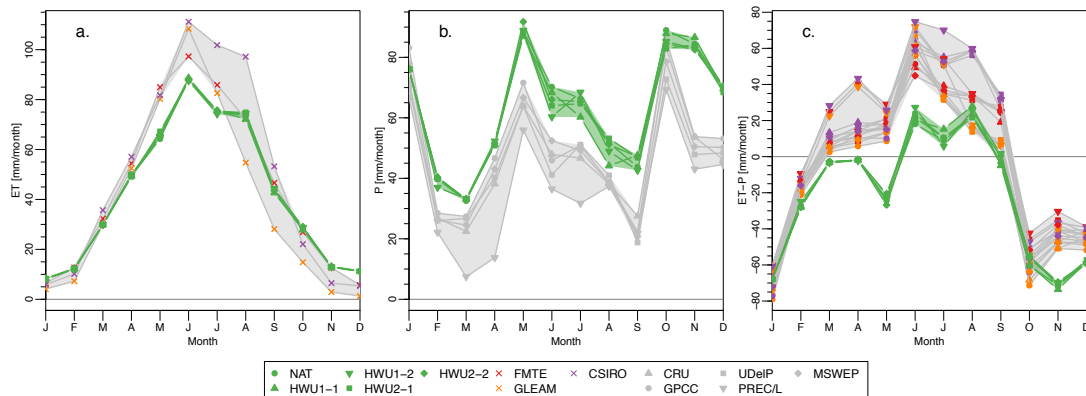
Figure E.1: Monthly time series of (a) evapotranspiration (ET), (b) precipitation (P) and (c) $ET - P$ in [$mm/month$] averaged over the Seine watershed for all gridded observational data sets and the simulations. The grey and green shaded areas indicate the maximum range covered by the observations and the simulations, respectively.
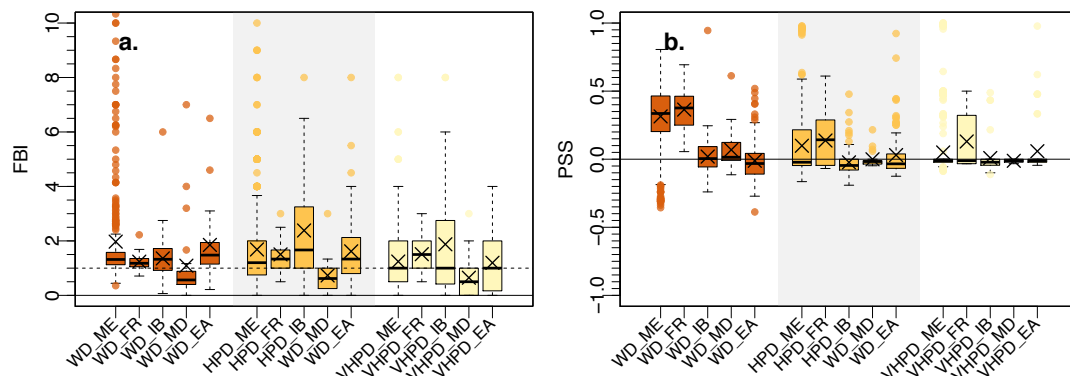
Figure E.2: (a) Frequency bias (FBI) and (b) Peirce's skill score (PSS) of the NAT simulation for wet days (WD, brown), heavy precipitation days (HPD, orange), and very heavy precipitation days (VHPD, yellow) over summer 2003. The boxplots show the distribution of each score over 696, 30, 76, 25 and 104 stations in the regions Mid-Europe (ME), France (FR), Iberian Peninsula (IB), Mediterranean (MD) and Eastern Europe (EA), respectively. The crosses denote the mean score over these stations.
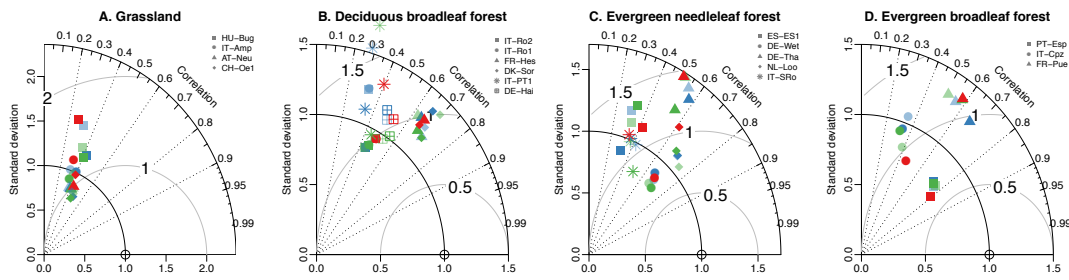
Figure E.3: Taylor diagrams of daily $ET - P$ for all (a) grassland, (b) deciduous broadleaf forest, (c) evergreen needleleaf forest and (d) evergreen broadleaf forest FLUXNET sites over the full year, using a normalized standard deviation. The colors indicate the simulation scenario (red: NAT, blues: HWU1, greens: HWU2) and the symbols the FLUXNET station.
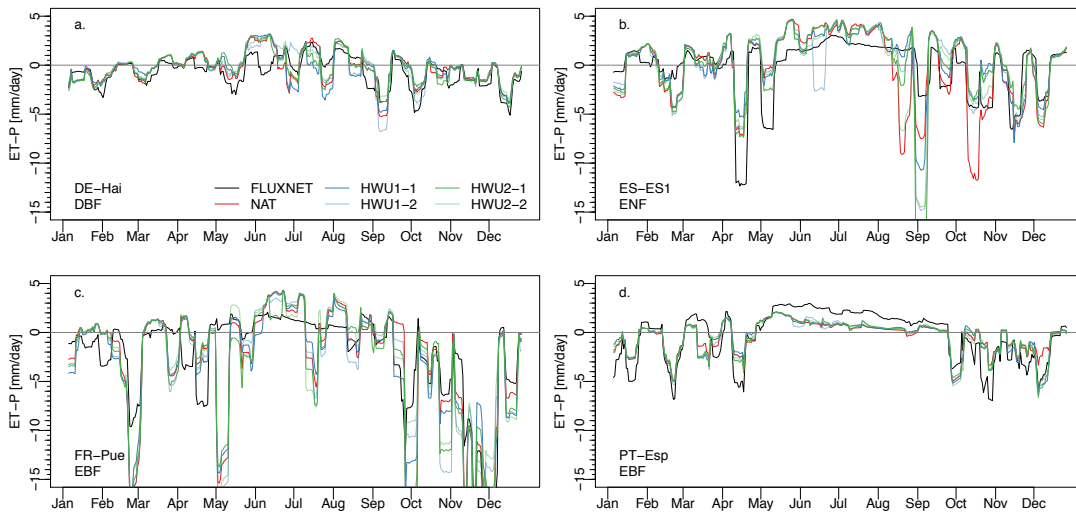
Figure E.4: Time series of observed and simulated daily $ET - P$ [mm/day] over 4 selected FLUXNET stations: (a) Hainich in Germany (deciduous broadleaf forest), (b) El Salar in Spain (evergreen needleleaf forest), (c) Puechabon in France (evergreen broadleaf forest) and (d) Espirra in Portugal (evergreen broadleaf forest). For better visualization, a 10-day moving average is plotted.
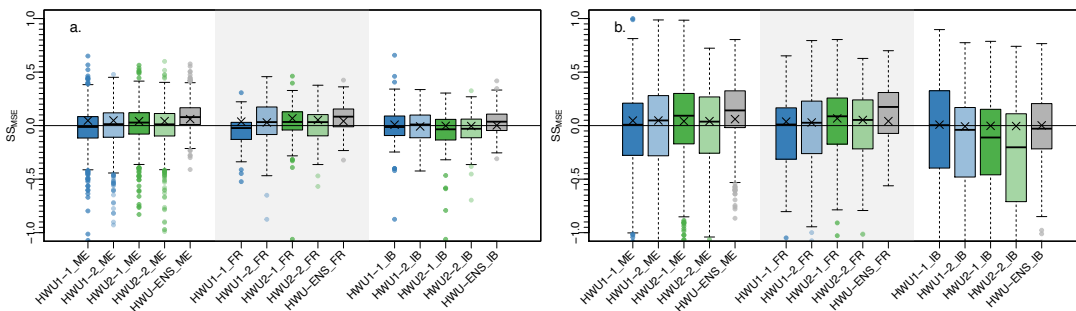
Figure E.5: MSE skill scores for improvements of the water use scenarios over the natural reference run for daily precipitation values for (a) the full year and (b) summer 2003. The boxplots show the distribution over 696, 30 and 76 stations in the regions Mid-Europe (ME), France (FR) and Iberian Peninsula (IB), respectively. The crosses indicate the mean MSE skill score over all stations for each water use scenario and the water use ensemble mean.
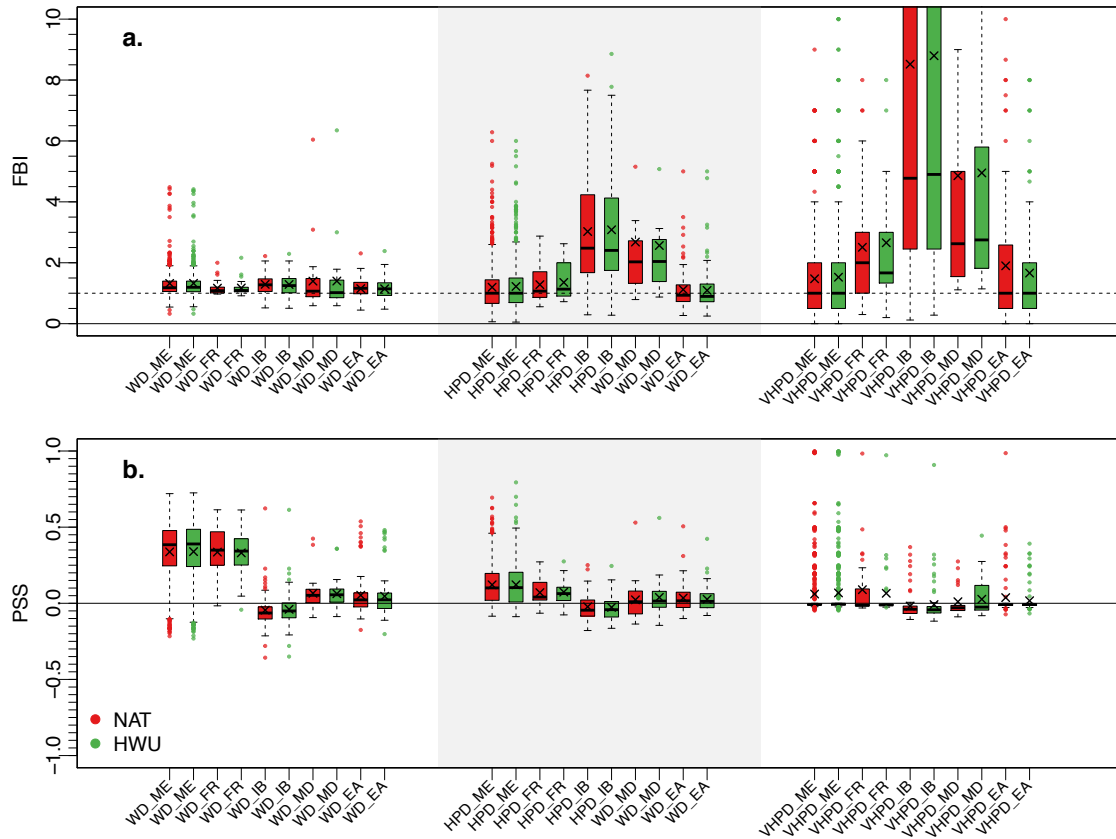
Figure E.6: (a) Frequency bias (FBI) and (b) Peirce's skill score (PSS) of the NAT (red) and HWU-ENS (green) simulation for wet days (WD), heavy precipitation days (HPD), and very heavy precipitation days (VHPD) over the full year 2003. The boxplots show the distribution of each score over the stations in the regions Mid-Europe (ME), France (FR), Iberian Peninsula (IB), Mediterranean (MD) and Eastern Europe (EA), respectively. The crosses denote the mean score over these stations.

Appendix F

Simulation output

Name	Unit	Description	2D	3D
ParFlow				
ψ	m	pressure head		x
S_W	-	relative saturation		x
CLM				
$FCEV$	W/m^2	canopy evaporation	x	
$FCTR$	W/m^2	canopy transpiration	x	
$FGEV$	W/m^2	ground evaporation	x	
$FIRA$	W/m^2	net infrared (longwave) radiation	x	
$FIRE$	W/m^2	emitted infrared (longwave) radiation	x	
FSA	W/m^2	absorbed solar radiation	x	
FSH	W/m^2	sensible heat	x	
FSH_G	W/m^2	sensible heat from ground	x	
FSH_V	W/m^2	sensible heat from vegetation	x	
FSM	W/m^2	snow melt heat flux	x	
FSH	W/m^2	sensible heat from ground	x	
FSR	W/m^2	reflected solar radiation	x	
$H2OCAN$	mm	intercepted water	x	
$H2OSNO$	mm	snow depth (liquid water)	x	
$H2OSOI$	mm^3/mm^3	volumetric soil water content		x
$Q2M$	kg/kg	2m specific humidity	x	
$SABV$	W/m^2	solar radiation absorbed by vegetation	x	
$SABG$	W/m^2	solar radiation absorbed by ground	x	
$SNOWAGE$	-	snow age	x	
$SNOWDP$	m	snow depth	x	
$SNOWICE$	kg/m^2	snow ice	x	
$SNOWLIQ$	kg/m^2	snow liquid water	x	
$SOILICE$	kg/m^2	soil ice	x	
$SOILLIQ$	kg/m^2	soil liquid water	x	
$TAUX$	$kg/m/s^2$	zonal surface stress	x	
$TAUY$	$kg/m/s^2$	meridional surface stress	x	
TG	K	ground temperature	x	
$TLAKE$	K	lake temperature	x	
TSA	K	2m air temperature	x	

Name	Unit	Description	2D	3D
TSNOW	*K*	snow temperature	x	
TSOI	*K*	soil temperature		x
TV	*K*	vegetation temperature	x	
TSNOW	*K*	snow temperature	x	
COSMO				
U	*m/s*	eastward wind component velocity		x
V	*m/s*	northward wind component velocity		x
W	*m/s*	vertical wind component velocity		x
T	*K*	air temperature		x
P	*Pa*	air pressure		x
QV	*kg/kg*	specific humidity		x
QC	*kg/kg*	specific cloud liquid water content		x
QI	*kg/kg*	specific cloud ice content		x
QS	*kg/kg*	specific snow content		x
QR	*kg/kg*	specific rain content		x
RELHUM	*%*	relative humidity		x
TKE	m^2/s^2	turbulent kinetic energy	x	
CLC	-	cloud area fraction		x
CLCT	-	total cloud area fraction	x	
CLC_CON	-	convective cloud area fraction	x	
CLW_CON	-	mass fraction of convective cloud liquid water in air	x	
FIS	m^2/s^2	surface geopotential	x	
PMSL	*Pa*	mean sea level pressure	x	
PS	*Pa*	surface pressure	x	
U_10M	*m/s*	U-component of 10m wind	x	
V_10M	*m/s*	V-component of 10m wind	x	
TQV	kg/m^2	vertically integrated precipitable water	x	
TQR	kg/m^2	vertically integrated rain water content	x	
TQC	kg/m^2	vertically integrated cloud water content	x	
TQS	kg/m^2	vertically integrated snow water content	x	
TQI	kg/m^2	vertically integrated cloud ice content	x	
TWATER	kg/m^2	vertically integrated total water content	x	
RAIN_GSP	kg/m^2	large scale rainfall	x	
RAIN_CON	kg/m^2	convective precipitation amount	x	
TOT_PREC	kg/m^2	total precipitation amount	x	
TCM	-	surface drag coefficient for momentum in air	x	

Name	Unit	Description	2D	3D
TCH	-	surface drag coefficient for heat in air	x	
HPBL	*m*	Height of the planetary boundary layer	x	
CEILING	*m*	Cloud ceiling height (above mean sea level)	x	
CAPE_ML	*J/kg*	convective available potential energy of a mean surface layer parcel	x	
CAPE_MU	*J/kg*	convective available potential energy of the most unstable parcel	x	
CIN_ML	*J/kg*	convective inhibition of a mean surface layer parcel	x	
CIN_MU	*J/kg*	convective inhibition of the most unstable parcel	x	
SNOW_GSP	kg/m^2	large scale snowfall	x	
SNOW_CON	kg/m^2	convective snowfall	x	
ASOB_S	W/m^2	averaged surface net downward shortwave radiation	x	
ATHB_S	W/m^2	averaged surface net downward longwave radiation	x	

Table F.1: Overview of the hourly model output variables from ParFlow, CLM and COSMO over the European CORDEX domain.

List of Abbreviations

ABL	Atmospheric Boundary Layer
ANOVA	Analysis of Variance
CAI	Climatologically Aided Interpolation
CAPE	Convective Available Potential Energy
CAMS	Climate Anomaly Monitoring System
CLM	Community Land Model
COSMO	Consortium for Small-Scale Modeling
CORDEX	Coordinated Regional Downscaling Experiment
CRU	Climate Research Unit
CSIRO	Commonwealth Scientific and Industrial Research Organization
DEM	Digital Elevation Model
DS	Downscaled/ Downscaling
DWD	Deutscher Wetterdienst
ECA&D	European Climate Assessment & Dataset
ECMWF	European Center for Medium-Range Weather Forecasts
EL	Equilibrium level
ENS	Ensemble
ESM	Earth System Modeling/ Earth System Models
FAO	Food and Agriculture Organization of the United Nations
FBI	Frequency Bias
FD	Free-Drainage boundary condition
FLUXNET	Flux Network
GCM	Global Climate Model
GHCN2	Global Historical Climatology Network
GLEAM	Global Land Evaporation Amsterdam Model
GPCC	Global Precipitation Climatology Centre
H-TESSEL	Tiled ECMWF Scheme for Surface Exchanges over Land with a revised hydrology
HFD	Hydro-facies distribution
HPC	High-Performance Computing
HPD	Heavy Precipitation Days
HWU	Human Water Use
JSC	Jülich Supercomputing Centre
LAI	Leaf Area Index
LSM	Land Surface Model
LCL	Lifting Condensation Level
LFC	Level of Free Convection
MCT	Model Coupling Toolkit
MHTOP	Monthly top height of a PFT
MHBOT	Monthly bottom height of a PFT
MODIS	Moderate Resolution Imaging Spectraradiometer
MSE	Mean Square Error
MSWEP	Multi-Source Weighted-Ensemble Precipitation
MTE	Model Tree Ensemble
M-O	Monin-Obukhov
NCAR	National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
OASIS	Ocean Atmosphere Sea Ice Soil coupler
PFT	Plant functional type

PSS	Peirce's Skill Score
RCM	Regional Climate Model
RMSD	Root Mean Square Difference
SAI	Stem Area Index
SM	Soil Moisture
SS	Skill Score
TerrSysMP	Terrestrial Systems Modeling Platform
TR32	Transregional Collaborative Research Centre 32
UDelP	University of Delaware – Precipitation
USGS	United States Geological Survey
VHPD	Very Heavy Precipitation Days
WD	Wet Days
WTD	Water Table Depth

List of Symbols

u	west-east wind component
v	north-south wind component
w	vertical wind component
T	temperature
p'	pressure deviation from reference state
p_0	reference pressure
p	pressure
P	precipitation
q_v	specific humidity
q_l	specific cloud water content
q_f	specific cloud ice content
q_r	specific water content of rain
q_{sn}	specific water content of snow
ρ_a	air density
T_a	lowest-level air temperature
T_a^*	height-corrected lowest-level air temperature
T_S	surface temperature
$q_{v,a}$	lowest-level specific humidity
$q_{v,s}$	specific humidity at the surface
R_{net}	net radiation
G	ground heat flux
H	sensible heat flux
LE	latent heat flux
ET	evapotranspiration
E_C	canopy evaporation
E_G	ground evaporation
T_C	canopy transpiration
LW_{up}	upward longwave radiation
SW	shortwave radiation
LW_{dw}	downward longwave radiation
α_S	albedo
q_{rain}	precipitation / specific water content of rain
q_e	evapotranspiration / specific water content of evapotranspiration
S_S	specific storage coefficient (ParFlow)
θ	soil moisture
ψ_p	pressure head
ψ_s	ponding depth
ϕ	porosity
q_S	source-sink term in 3D Richards equation (ParFlow)
$k(x)$	saturated hydraulic conductivity
$k_r(\psi_p)$	relative permeability
z	depth
q	water flux
t	time
S_W	relative saturation
v_x	flow velocity in x-direction
v_y	flow velocity in y-direction
n_m	Manning's coefficient
q_{out}	outflow at an outlet

S_f	friction slope
Ψ	prognostic variable
Ψ^*	outer state of a variable (Davies boundary)
D_x	domain size in x-direction
D_y	domain size in y-direction
m	meridional wave number
n	zonal wave number
k_m	x-direction wave vector component
k_n	y-direction wave vector component
$A_{m,n}$	spectral Fourier coefficient of the regional (interior) model
$A_{m,n}^o$	spectral Fourier coefficient of the global (outer) model
α	spectral nudging coefficient
$L(\Psi)$	original model solution
γ	nudging coefficient
λ	latent heat of vaporization
R_d	specific gas constant for dry air
c_p	specific heat capacity for dry air
ρ	Pearson correlation coefficient
C_{SI}	Continental sink for atmospheric water
C_{SO}	Continental source for atmospheric water
R	runoff
S	subsurface water storage
$div(Q_g)$	groundwater divergence
$div(Q)$	atmospheric divergence of specific humidity
Y	random variable
a	linear regression coefficient
μ	mean
σ	standard deviation
$\mathcal{N}$	normal distribution
H_0	null hypothesis
SST	total sum of squares (ANOVA)
SSA	treatment sum of squares (ANOVA)
SSE	sum of squared errors (ANOVA)
R^2	(adjusted) coefficient of multiple determination (ANOVA)
SS_{MSE}	MSE skill score
z_{EL}	height of the equilibrium level
z_{LFC}	height of the level of free convection
T_v	virtual temperature
g	gravity acceleration

List of Figures

List of Tables

Bibliography

Addor, N., M. Rohrer, R. Furrer, and J. Seibert (2016). "Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment". *Journal of Geophysical Research: Atmospheres*, 121 (5), 2075–2089.

Allen, M. R. and W. J. Ingram (2002). "Constraints on future changes in climate and the hydrologic cycle". *Nature*, 419 (6903), 224–232.

Alter, R. E., Y. Fan, B. R. Lintner, C. P. Weaver, R. E. Alter, Y. Fan, B. R. Lintner, and C. P. Weaver (2015). "Observational Evidence that Great Plains Irrigation Has Enhanced Summer Precipitation Intensity and Totals in the Midwestern United States". *Journal of Hydrometeorology*, 16 (4), 1717–1735.

Anyah, R. O., C. P. Weaver, G. Miguez-Macho, Y. Fan, and A. Robock (2008). "Incorporating water table dynamics in climate modeling: 3. Simulated groundwater influence on coupled land-atmosphere variability". *Journal of Geophysical Research*, 113 (D7).

Ashby, S. F. and R. D. Falgout (1996). "A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations". *Nuclear Science and Engineering*, 124 (1), 145–159.

Baatz, D., W. Kurtz, H.-J. Hendricks-Franssen, H. Vereecken, and S. J. Kollet (2017). "Catchment tomography – An approach for spatial parameter estimation". *Advances in Water Resources*, 107, 147–159.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt (2011). "Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities". *Monthly Weather Review*, 139 (12), 3887–3905.

Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R. Evans, J. Fuentes, A. Goldstein, G. Katul, B. Law, X. Lee, Y. Malhi, T. Meyers, W. Munger, W. Oechel, K. T. Paw U, K. Pilegaard, H. P. Schmid, R. Valentini, S. Verma, T. Vesala, K. Wilson, and S. Wofsy (2001). "FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities". *Bulletin of the American Meteorological Society*, 82 (11), 2415–2434.

Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, E. Brun, H. Cloke, D. Dee, E. Dutra, J. Muñoz-Sabater, F. Pappenberger, P. de Rosnay, T. Stockdale, and F. Vitart (2015). "ERA-Interim/Land: a global land surface reanalysis data set". *Hydrology and Earth System Sciences*, 19 (1), 389–407.

Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. Betts (2009). "A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System". *Journal of Hydrometeorology*, 10 (3), 623–643.

Barlage, M., M. Tewari, F. Chen, G. Miguez-Macho, Z.-L. Yang, and G.-Y. Niu (2015). "The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP". *Climatic Change*, 129 (3), 485–498.

Barnston, A. G. and P. T. Schickendanz (1984). "The effect of irrigation on warm season precipitation in the southern Great Plains". *Journal of Climate and Applied Meteorology*, 23 (6), 865–888.

Bauer, P., A. Thorpe, and G. Brunet (2015). "The quiet revolution of numerical weather prediction". *Nature*, 525 (7567), 47–55.

Beck, H. E., N. Vergopolan, M. Pan, V. Levizzani, A. I.J. M. van Dijk, G. P. Weedon, L. Brocca, F. Pappenberger, G. J. Huffman, and E. F. Wood (2017a). "Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling". *Hydrology and Earth System Sciences*, 21 (12), 6201–6217.

Beck, H. E., A. I.J. M. van Dijk, V. Levizzani, J. Schellekens, D. G. Miralles, B. Martens, and A. de Roo (2017b). "MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data". *Hydrology and Earth System Sciences*, 21 (1), 589–615.

Beniston, M. (2004). "The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations". *Geophysical Research Letters*, 31 (2).

Beniston, M. and D. B. Stephenson (2004). "Extreme climatic events and their evolution under changing climatic conditions". *Global and Planetary Change*, 44 (1-4), 1–9.

Betts, A. K., J. H. Ball, A. Beljaars, M. J. Miller, and P. A. Viterbo (1996). "The land surface-atmosphere interaction: A review based on observational and global modeling perspectives". *Journal of Geophysical Research: Atmospheres*, 101 (D3), 7209–7225.

Biemans, H., R. W. A. Hutjes, P. Kabat, B. J. Strengers, D. Gerten, and S. Rost (2009). "Effects of Precipitation Uncertainty on Discharge Calculations for Main River Basins". *Journal of Hydrometeorology*, 10 (4), 1011–1025.

Bierkens, M. F. P. (2015). "Global hydrology 2015: State, trends, and directions". *Water Resources Research*, 51 (7), 4923–4947.

Bierkens, M. F. P., V. A. Bell, P. Burek, N. Chaney, L. E. Condon, C. H. David, A. de Roo, P. Döll, N. Drost, J. S. Famiglietti, M. Flörke, D. J. Gochis, P. Houser, R. Hut, J. Keune, S. J. Kollet, R. M. Maxwell, J. T. Reager, L. Samaniego, E. Sudicky, E. H. Sutanudjaja, N. van de Giesen, H. Winsemius, and E. F. Wood (2015). "Hyper-resolution global hydrological modelling: What is next? Everywhere and locally relevant". *Hydrological Processes*, 29 (2), 310–320.

Black, E., M. Blackburn, G. Harrison, B. J. Hoskins, and J. Methven (2004). "Factors contributing to the summer 2003 European heatwave". *Weather*, 59 (8), 217–223.

Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson (2002). "Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models". *Global Biogeochemical Cycles*, 16 (2), 5–1–5–23.

Bonan, G. B., P. J. Lawrence, K. W. Oleson, S. Levis, M. Jung, M. Reichstein, D. M. Lawrence, and S. C. Swenson (2011). "Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data". *Journal of Geophysical Research: Biogeosciences*, 116 (G2).

Bosshard, T., M. Carambia, K. Goergen, S. Kotlarski, P. Krahe, M. Zappa, and C. Schär (2013). "Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections". *Water Resources Research*, 49 (3), 1523–1536.

Brubaker, K. L., D. Entekhabi, and P. S. Eagleson (1993). "Estimation of Continental Precipitation Recycling". *Journal of Climate*, 6 (6), 1077–1089.

Brutsaert, W. (2005). *Hydrology: an introduction.* Cambridge University Press.

Butts, M. B., M. Drews, M. A. Larsen, S. Lerer, S. H. Rasmussen, J. Grooss, J. Overgaard, J. C. Refsgaard, O. B. Christensen, and J. H. Christensen (2014). "Embedding complex hydrology in the regional climate system – Dynamic coupling across different modelling domains". *Advances in Water Resources*, 74, 166–184.

Campoy, A., A. Ducharne, F. Cheruy, F. Hourdin, J. Polcher, and J. C. Dupont (2013). "Response of land surface fluxes and precipitation to different soil bottom hydrological conditions in a general circulation model". *Journal of Geophysical Research: Atmospheres*, 118 (19), 10725–10739.

Casanueva, A., S. Kotlarski, S. Herrera, J. Fernández, J. M. Gutiérrez, F. Boberg, A. Colette, O. B. Christensen, K. Goergen, D. Jacob, K. Keuler, G. Nikulin, C. Teichmann, and R. Vautard (2016). "Daily precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected high-resolution simulations". *Climate Dynamics*, 47 (3-4), 719–737.

Chen, F. and J. Dudhia (2001). "Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity". *Monthly Weather Review*, 129 (4), 569–585.

Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin (2002). "Global land precipitation: A 50-yr monthly analysis based on gauge observations". *Journal of Hydrometeorology*, 3 (3), 249–266.

Chen, X. and Q. Hu (2004). "Groundwater influences on soil moisture and surface evaporation". *Journal of Hydrology*, 297 (1-4), 285–300.

Christensen, J. H., T. R. Carter, M. Rummukainen, and G. Amanatidis (2007). "Evaluating the performance and utility of regional climate models: The PRUDENCE project". *Climatic Change*, 81, 1–6.

Clark, M. P., Y. Fan, D. M. Lawrence, J. C. Adam, D. Bolster, D. J. Gochis, R. P. Hooper, M. Kumar, L. R. Leung, D. S. Mackay, R. M. Maxwell, C. Shen, S. C. Swenson, and X. Zeng (2015). "Improving the representation of hydrologic processes in Earth System Models". *Water Resources Research*, 51 (8), 5929–5956.

Condon, L. E. and R. M. Maxwell (2013). "Implementation of a linear optimization water allocation algorithm into a fully integrated physical hydrology model". *Advances in Water Resources*, 60, 135–147.

Condon, L. E. and R. M. Maxwell (2014a). "Feedbacks between managed irrigation and water availability: Diagnosing temporal and spatial patterns using an integrated hydrologic model". *Water Resources Research*, 50 (3), 2600–2616.

— (2014b). "Groundwater-fed irrigation impacts spatially distributed temporal scaling behavior of the natural system: a spatio-temporal framework for understanding water management impacts". *Environmental Research Letters*, 9 (3), 034009.

— (2015). "Evaluating the relationship between topography and groundwater using outputs from a continental-scale integrated hydrology model". *Water Resources Research*, 51 (8), 6602–6621.

Cook, B. I., S. P. Shukla, M. J. Puma, and L. S. Nazarenko (2015). "Irrigation as an historical climate forcing". *Climate Dynamics*, 44 (5-6), 1715–1730.

Daniel, B. (2015). "Big Data and analytics in higher education: Opportunities and challenges". *British Journal of Educational Technology*, 46 (5), 904–920.

Davies, H. C. (1976). "A laterul boundary formulation for multi-level prediction models". *Quarterly Journal of the Royal Meteorological Society*, 102 (432), 405–418.

Davin, E. L., E. Maisonnave, and S. I. Seneviratne (2016). "Is land surface processes representation a possible weak link in current Regional Climate Models?" *Environmental Research Letters*, 11 (7), 074027.

Davison, J. H., H.-T. Hwang, E. A. Sudicky, and J. C. Lin (2015). "Coupled atmospheric, land surface, and subsurface modeling: Exploring water and energy feedbacks in three-dimensions". *Advances in Water Resources*, 86, 73–85.

Davison, J. H., H. Hwang, E. A. Sudicky, D. V. Mallia, and J. C. Lin (2018). "Full Coupling Between the Atmosphere, Surface, and Subsurface for Integrated Hydrologic Simulation". *Journal of Advances in Modeling Earth Systems*, 10 (1), 43–53.

de Vrese, P., S. Hagemann, and M. Claussen (2016). "Asian irrigation, African rain: Remote impacts of irrigation". *Geophysical Research Letters*, 43 (8), 3737–3745.

DeAngelis, A., F. Dominguez, Y. Fan, A. Robock, M. D. Kustu, and D. Robinson (2010). "Evidence of enhanced precipitation due to irrigation over the Great Plains of the United States". *Journal of Geophysical Research: Atmospheres*, 115 (15), D15115.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011). "The ERA-Interim reanalysis: configuration and performance of the data assimilation system". *Quarterly Journal of the Royal Meteorological Society*, 137 (656), 553–597.

Della-Marta, P. M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner (2007). "Summer heat waves over western Europe 1880 - 2003, their relationship to large-scale forcings and predictability". *Climate Dynamics*, 29 (2-3), 251–275.

Di Luca, A., R. de Elía, and R. Laprise (2015). "Challenges in the quest for added value of regional climate dynamical downscaling". *Current Climate Change Reports*, 1 (1), 10–21.

Dirmeyer, P. A. (2011). "The terrestrial segment of soil moisture-climate coupling". *Geophysical Research Letters*, 38 (16), L16702.

Dirmeyer, P. A., R. D. Koster, and Z. Guo (2006). "Do global models properly represent the feedback between land and atmosphere?" *Journal of Hydrometeorology*, 7 (6), 1177–1198.

Dirmeyer, P. A., Z. Wang, M. J. Mbuh, and H. E. Norton (2014). "Intensified land surface control on boundary layer growth in a changing climate". *Geophysical Research Letters*, 41 (4), 1290–1294.

Döll, P., F. Kaspar, and B. Lehner (2003). "A global hydrological model for deriving water availability indicators: model tuning and validation". *Journal of Hydrology*, 270 (1-2), 105–134.

Doms, G. and U. Schättler (2002). *A description of the nonhydrostatic regional model LM. Part I: Dynamics and Numerics*. Tech. Report. Deutscher Wetterdienst, Offenbach am Main.

Doms, G., J. Förstner, E. Heise, H.-J. Herzog, D. Mironov, M. Raschendorfer, T. Reinhardt, B. Ritter, R. Schrodin, J.-P. Schulz, and G. Vogel (2011). *A description of the nonhydrostatic regional COSMO model. Part II: physical parameterization*. Tech. Report. Deutscher Wetterdienst, Offenbach am Main.

Doswell III, C. A. and E. N. Rasmussen (1994). "The effect of neglecting the virtual temperature correction on CAPE calculations". *Weather and Forecasting*, 9 (4), 625–629.

Douglas, E. M., A. Beltrán-Przekurat, D. Niyogi, R. A. Pielke Sr., and C. J. Vörösmarty (2009). "The impact of agricultural intensification and irrigation on land–atmosphere interactions and Indian monsoon precipitation?A mesoscale modeling perspective". *Global and Planetary Change*, 67 (1-2), 117–128.

Douville, H. (2004). "Relevance of soil moisture for seasonal atmospheric predictions: is it an initial value problem?" *Climate Dynamics*, 22 (4), 429–446.

Eltahir, E. A. (1998). "A soil moisture–rainfall feedback mechanism: 1. Theory and observations". *Water Resources Research*, 34 (4), 765–776.

Famiglietti, J. S. (2014). "The global groundwater crisis". *Nature Climate Change*, 4 (11), 945–948.

Fan, Y. (2015). "Groundwater in the Earth's critical zone: Relevance to large-scale patterns and processes". *Water Resources Research*, 51 (5), 3052–3069.

Fan, Y., H. Li, and G. Miguez-Macho (2013). "Global patterns of groundwater table depth." *Science*, 339 (6122), 940–943.

Fan, Y., G. Miguez-Macho, E. G. Jobbágy, R. B. Jackson, and C. Otero-Casal (2017). "Hydrologic regulation of plant rooting depth". *Proceedings of the National Academy of Sciences of the United States of America*, 201712381. DOI: 10.1073/pnas.1712381114.

FAO (1988). "UNESCO soil map of the world, revised legend". *World Resources Report*, 60, 138.

Fennessy, M. J. and J. Kinter III (2011). "Climatic feedbacks during the 2003 European heat wave". *Journal of Climate*, 24 (23), 5953–5967.

Ferguson, I. M. and R. M. Maxwell (2010). "Role of groundwater in watershed response and land surface feedbacks under climate change". *Water Resources Research*, 46 (10), W00F02.

Ferguson, I. M. and R. M. Maxwell (2012). "Human impacts on terrestrial hydrology: climate change versus pumping and irrigation". *Environmental Research Letters*, 7 (4), 044022.

Ferranti, L. and P. A. Viterbo (2006). "The European Summer of 2003: Sensitivity to Soil Water Initial Conditions". *Journal of Climate*, 19 (15), 3659–3680.

Feser, F. (2006). "Enhanced detectability of added value in limited-area model results separated into different spatial scales". *Monthly Weather Review*, 134 (8), 2180–2190.

Findell, K. L., P. Gentine, B. R. Lintner, and C. Kerr (2011). "Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation". *Nature Geoscience*, 4 (7), 434–439.

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

Foken, T. (2006). "50 years of the Monin–Obukhov similarity theory". *Boundary-Layer Meteorology*, 119 (3), 431–447.

Friedl, M. A., D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf (2002). "Global land cover mapping from MODIS: algorithms and early results". *Remote Sensing of Environment*, 83 (1-2), 287–302.

Gasper, F., K. Goergen, P. Shrestha, M. Sulis, J. Rihani, M. Geimer, and S. J. Kollet (2014). "Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP v1.0) in a massively parallel supercomputing environment - a case study on JUQUEEN (IBM Blue Gene/Q)". *Geoscientific Model Development*, 7 (5), 2531–2543.

Gilbert, J. M., R. M. Maxwell, and D. J. Gochis (2017). "Effects of Water-Table Configuration on the Planetary Boundary Layer over the San Joaquin River Watershed, California". *Journal of Hydrometeorology*, 18 (5), 1471–1488.

Gimeno, L., A. Stohl, R. M. Trigo, F. Dominguez, K. Yoshimura, L. Yu, A. Drumond, A. M. Durán-Quesada, and R. Nieto (2012). "Oceanic and terrestrial sources of continental precipitation". *Reviews of Geophysics*, 50, 1–41.

Giorgi, F. (1995). "Perspectives for regional earth system modeling". *Global and Planetary Change*, 10 (1-4), 23–42.

Giorgi, F., C. Jones, and G. R. Asrar (2009). "Addressing climate information needs at the regional level: the CORDEX framework". *World Meteorological Organization (WMO) Bulletin*, 58 (3), 175–183.

Gleeson, T., L. Marklund, L. Smith, and A. H. Manning (2011a). "Classifying the water table at regional to continental scales". *Geophysical Research Letters*, 38 (5), L05401.

Gleeson, T., L. Smith, N. Moosdorf, J. Hartmann, H. H. Dürr, A. H. Manning, L. P. H. van Beek, and A. M. Jellinek (2011b). "Mapping permeability over the surface of the Earth". *Geophysical Research Letters*, 38 (2), L02401.

Greve, P., B. Orlowsky, B. Mueller, J. Sheffield, M. Reichstein, and S. I. Seneviratne (2014). "Global assessment of trends in wetting and drying over land". *Nature Geoscience*, 7 (10), 716–721.

Guillod, B. P., B. Orlowsky, D. G. Miralles, A. J. Teuling, P. D. Blanken, N. Buchmann, P. Ciais, M. Ek, K. L. Findell, P. Gentine, B. R. Lintner, R. L. Scott, B. van den Hurk, and S. I. Seneviratne (2014). "Land-surface controls on afternoon precipitation diagnosed from observational data: uncertainties and confounding factors". *Atmospheric Chemistry and Physics*, 14 (16), 8343–8367.

Guillod, B. P., B. Orlowsky, D. G. Miralles, A. J. Teuling, and S. I. Seneviratne (2015). "Reconciling spatial and temporal soil moisture effects on afternoon rainfall". *Nature Communications*, 6, 6443.

Haddeland, I., D. P. Lettenmaier, and T. Skaugen (2006a). "Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins". *Journal of Hydrology*, 324 (1-4), 210–223.

Haddeland, I., T. Skaugen, and D. P. Lettenmaier (2006b). "Anthropogenic impacts on continental surface water fluxes". *Geophysical Research Letters*, 33 (8), L08406.

Harding, K. J. and P. K. Snyder (2012). "Modeling the atmospheric response to irrigation in the Great Plains. Part II: The precipitation of irrigated water and changes in precipitation recycling". *Journal of Hydrometeorology*, 13 (6), 1687–1703.

Harris, I., P. D. Jones, T. Osborn, and D. H. Lister (2014). "Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset". *International Journal of Climatology*, 34 (3), 623–642.

Hashem, I. A. T., I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. U. Khan (2015). "The rise of "big data" on cloud computing: Review and open research issues". *Information Systems*, 47, 98–115.

Hazenberg, P., P. Broxton, D. Gochis, G.-Y. Niu, L. A. Pangle, J. D. Pelletier, P. Troch, and X. Zeng (2016). "Testing the hybrid-3-D hillslope hydrological model in a controlled environment". *Water Resources Research*, 52 (2), 1089–1107.

He, X., Y. Wada, N. Wanders, and J. Sheffield (2017). "Human water management intensifies hydrological drought in California". *Geophysical Research Letters*, 44, 1777–1785.

Hirsch, A. L., M. Wilhelm, E. L. Davin, W. Thiery, and S. I. Seneviratne (2017). "Can climate-effective land management reduce regional warming?" *Journal of Geophysical Research: Atmospheres*, 122 (4), 2269–2288.

Hirschi, M., S. I. Seneviratne, V. Alexandrov, F. Boberg, C. Boroneant, O. Christensen, H. Formayer, B. Orlowsky, and P. Stepanek (2010). "Observational evidence for soil-moisture impact on hot extremes in southeastern Europe". *Nature Geoscience*, 4 (1), 17–21.

Hoekstra, A. Y. and M. M. Mekonnen (2012). "The water footprint of humanity". *Proceedings of the National Academy of Sciences of the United States of America*, 109 (9), 3232–3237.

Hohenegger, C. and C. Schär (2007). "Atmospheric predictability at synoptic versus cloud-resolving scales". *Bulletin of the American Meteorological Society*, 88 (11), 1783–1794.

Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schär (2009). "The soil moisture–precipitation feedback in simulations with explicit and parameterized convection". *Journal of Climate*, 22 (19), 5003–5020.

Huntington, T. G. (2006). "Evidence for intensification of the global water cycle: Review and synthesis". *Journal of Hydrology*, 319 (1), 83–95.

Jaeger, E. B. and S. I. Seneviratne (2010). "Impact of soil moisture–atmosphere coupling on European climate extremes and trends in a regional climate model". *Climate Dynamics*, 36 (9-10), 1919–1939.

Jaramillo, F. and G. Destouni (2015). "Local flow regulation and irrigation raise global human water consumption and footprint". *Science*, 350 (6265), 1248–1251.

Ji, P., X. Yuan, and X.-Z. Liang (2017). "Do Lateral Flows Matter for the Hyperresolution Land Surface Modeling?" *Journal of Geophysical Research: Atmospheres*, 122 (22), 12077–12092.

Johnson, K. D., D. Entekhabi, and P. S. Eagleson (1993). "The implementation and validation of improved land-surface hydrology in an atmospheric general circulation model". *Journal of Climate*, 6 (6), 1009–1026.

Jolliffe, I. T. and D. B. Stephenson (2012). *Forecast verification: a practitioner's guide in atmospheric science.* Oxford: Wiley.

Jones, J. and C. S. Woodward (2001). "Newton-Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems". *Advances in Water Resources*, 24, 763–774.

Jung, M., M. Reichstein, and A. Bondeau (2009). "Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model". *Biogeosciences*, 6 (10), 2001–2013.

Jung, M., M. Reichstein, P. Ciais, S. I. Seneviratne, J. Sheffield, M. L. Goulden, G. Bonan, A. Cescatti, J. Chen, R. A. M. de Jeu, A. J. Dolman, W. Eugster, D. Gerten, D. Gianelle, N. Gobron, J. Heinke, J. Kimball, B. E. Law, L. Montagnani, M. Qiaozhen, B. Mueller, K. Oleson, D. Papale, A. D. Richardson, O. Roupsard, S. Running, E. Tomelleri, N. Viovy, U. Weber, C. Williams, E. Wood, S. Zaehle, and K. Zhang (2010). "Recent decline in the global land evapotranspiration trend due to limited moisture supply". *Nature*, 467 (7318), 951.

Jung, M., M. Reichstein, H. A. Margolis, A. Cescatti, A. D. Richardson, M. A. Arain, A. Arneth, C. Bernhofer, D. Bonal, J. Chen, D. Gianelle, N. Gobron, G. Kiely, W. Kutsch, G. Lasslop, B. E. Law, A. Lindroth, L. Merbold, L. Montagnani, E. J. Moors, D. Papale, M. Sottocornola, F. Vaccari, and C. Williams (2011). "Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations". *Journal of Geophysical Research: Biogeosciences*, 116 (G3).

Karl, T. R. and K. E. Trenberth (2003). "Modern global climate change". *Science*, 302 (5651), 1719–1723.

Katragkou, E., M. García-Díez, R. Vautard, S. Sobolowski, P. Zanis, G. Alexandri, R. M. Cardoso, A. Colette, J. Fernandez, A. Gobiet, K. Goergen, T. Karacostas, S. Knist, S. Mayer, P. M. M. Soares, I. Pytharoulis, I. Tegoulias, A. Tsikerdekis, and D. Jacob

(2015). "Regional climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble". *Geoscientific Model Development*, 8 (3), 603–618.

Keune, J., F. Gasper, K. Goergen, A. Hense, P. Shrestha, M. Sulis, and S. J. Kollet (2016). "Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003". *Journal of Geophysical Research: Atmospheres*, 121 (22), 13301–13325.

Keune, J., M. Sulis, S. J. Kollet, S. Siebert, and Y. Wada (2018). "Human Water Use Impacts on the Strength of the Continental Sink for Atmospheric Water". *Geophysical Research Letters*, 45 (9), 4068–4076.

Keys, P. W., L. Wang-Erlandsson, and L. J. Gordon (2016). "Revealing invisible water: moisture recycling as an ecosystem service". *PloS one*, 11 (3), e0151993.

Keys, P. W., L. Wang-Erlandsson, L. J. Gordon, V. Galaz, and J. Ebbesson (2017). "Approaching moisture recycling governance". *Global Environmental Change*, 45, 15–23.

Klein Tank, A. M. G., J. B. Wijngaard, G. P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A. F. V. van Engelen, E. Forland, M. Mietus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnas, J. Antonio López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L. V. Alexander, and P. Petrovic (2002). "Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment". *International Journal of Climatology*, 22 (12), 1441–1453.

Koirala, S., M. Jung, M. Reichstein, I. E. M. de Graaf, G. Camps-Valls, K. Ichii, D. Papale, B. Ráduly, C. R. Schwalm, G. Tramontana, and N. Carvalhais (2017). "Global distribution of groundwater-vegetation spatial covariation". *Geophysical Research Letters*, 44 (9), 4134–4142.

Kollet, S. J. and R. M. Maxwell (2006). "Integrated surface - groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model". *Advances in Water Resources*, 29 (7), 945–958.

— (2008). "Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model". *Water Resources Research*, 44 (2), W02402.

Konikow, L. F. and E. Kendy (2005). "Groundwater depletion: A global problem". *Hydrogeology Journal*, 13 (1), 317–320.

Koster, R. D., M. J. Suarez, R. W. Higgins, and H. M. Van den Dool (2003). "Observational evidence that soil moisture variations affect precipitation". *Geophysical Research Letters*, 30 (5).

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada (2004). "Regions of strong coupling between soil moisture and precipitation". *Science*, 305 (5687), 1138–1140.

Kotlarski, S., K. Keuler, O. B. Christensen, A. Colette, M. Déqué, A. Gobiet, K. Goergen, D. Jacob, D. Lüthi, E. van Meijgaard, G. Nikulin, C. Schär, C. Teichmann, R. Vautard, K. Warrach-Sagi, and V. Wulfmeyer (2014). "Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble". *Geoscientific Model Development*, 7 (4), 1297–1333.

Kotlarski, S., P. Szabó, S. Herrera, O. Räty, K. Keuler, P. M. Soares, R. M. Cardoso, T. Bosshard, C. Pagé, F. Boberg, J. M. Gutiérrez, F. A. Isotta, A. Jaczewski, F. Kreienkamp, M. A. Liniger, C. Lussana, and K. Pianko-Kluczyńska (2017). "Observational uncertainty and regional climate model evaluation: a pan-European perspective". *International Journal of Climatology*. DOI: 10.1002/joc.5249.

Krakauer, N. Y., H. Li, and Y. Fan (2014). "Groundwater flow across spatial scales: importance for climate modeling". *Environmental Research Letters*, 9 (3), 034003.

Kueppers, L. M. and M. A. Snyder (2012). "Influence of irrigated agriculture on diurnal surface energy and water fluxes, surface climate, and atmospheric circulation in California". *Climate Dynamics*, 38 (5-6), 1017–1029.

Kueppers, L. M., M. A. Snyder, and L. C. Sloan (2007). "Irrigation cooling effect: Regional climate forcing by land-use change". *Geophysical Research Letters*, 34 (3), L03703.

Kueppers, L. M., M. A. Snyder, L. C. Sloan, D. Cayan, J. Jin, H. Kanamaru, M. Kanamitsu, N. L. Miller, M. Tyree, H. Du, and B. Weare (2008). "Seasonal temperature responses to land-use change in the western United States". *Global and Planetary Change*, 60 (3-4), 250–264.

Kurtz, W., G. He, S. J. Kollet, R. M. Maxwell, H. Vereecken, and H.-J. Hendricks-Franssen (2016). "TerrSysMP–PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface–subsurface model". *Geoscientific Model Development*, 9 (4), 1341–1360.

Lam, H. A., D. J. Karssenberg, B. J.J. M. van den Hurk, and M. F. P. Bierkens (2011). "Spatial and temporal connections in groundwater contribution to evaporation". *Hydrology and Earth System Sciences*, 15, 2621–2630.

Larsen, M. A. D., J. C. Refsgaard, M. Drews, M. B. Butts, K. H. Jensen, J. H. Christensen, and O. B. Christensen (2014). "Results from a full coupling of the HIRHAM regional climate model and the MIKE SHE hydrological model for a Danish catchment". *Hydrology and Earth System Sciences*, 18 (11), 4733–4749.

Lawston, P. M., J. A. Santanello, and S. V. Kumar (2017). "Irrigation Signals Detected From SMAP Soil Moisture Retrievals". *Geophysical Research Letters*, 44 (23), 11860–11867.

Legates, D. R. and C. J. Willmott (1990). "Mean seasonal and spatial variability in gauge-corrected, global precipitation". *International Journal of Climatology*, 10 (2), 111–127.

Lehner, B., K. Verdin, and A. Jarvis (2006). *HydroSHEDS technical documentation, version 1.0*. Tech. Report. World Wildlife Fund US, Washington, DC, 1–27.

Leng, G., M. Huang, Q. Tang, H. Gao, and L. R. Leung (2014). "Modeling the effects of groundwater-fed irrigation on terrestrial hydrology over the conterminous United States". *Journal of Hydrometeorology*, 15 (3), 957–972.

Leuning, R., Y. Zhang, A. Rajaud, H. Cleugh, and K. Tu (2008). "A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation". *Water Resources Research*, 44 (10), W10419.

Liu, X., Q. Tang, H. Cui, M. Mu, D. Gerten, S. N. Gosling, Y. Masaki, Y. Satoh, and Y. Wada (2017). "Multimodel uncertainty changes in simulated river flows induced by human impact parameterizations". *Environmental Research Letters*, 12 (2), 025009.

Lobell, D., G. Bala, A. Mirin, T. Phillips, R. M. Maxwell, and D. Rotman (2009). "Regional differences in the influence of irrigation on climate". *Journal of Climate*, 22 (8), 2248–2255.

Lobell, D. B., C. J. Bonfils, L. M. Kueppers, and M. A. Snyder (2008). "Irrigation cooling effect on temperature and heat index extremes". *Geophysical Research Letters*, 35 (9), L09705.

Long, D., B. R. Scanlon, L. Longuevergne, A. Y. Sun, D. N. Fernando, and H. Save (2013). "GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas". *Geophysical Research Letters*, 40 (13), 3395–3401.

Lu, Y. and L. M. Kueppers (2015). "Increased heat waves with loss of irrigation in the United States". *Environmental Research Letters*, 10 (6), 064010.

MacLeod, D. A., H. L. Cloke, F. Pappenberger, and A. Weisheimer (2016). "Improved seasonal prediction of the hot summer of 2003 over Europe through better representation of uncertainty in the land surface". *Quarterly Journal of the Royal Meteorological Society*, 142 (694), 79–90.

Martens, B., D. G. Miralles, H. Lievens, R. Van Der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. Dorigo, and N. Verhoest (2017). "GLEAM v3: Satellite-based land evaporation and root-zone soil moisture". *Geoscientific Model Development*, 10 (5), 1903–1925.

Maxwell, R. M., L. E. Condon, and S. J. Kollet (2014). "Simulation of groundwater and surface water over the continental US using a hyperresolution, integrated hydrologic model". *Geoscientific Model Development Discussions*, 7, 7317–7349.

Maxwell, R. M., L. E. Condon, and S. J. Kollet (2015). "A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3". *Geoscientific Model Development*, 8 (3), 923–937.

Maxwell, R. M. (2013). "A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling". *Advances in Water Resources*, 53, 109–117.

Maxwell, R. M., F. K. Chow, and S. J. Kollet (2007). "The groundwater – land-surface – atmosphere connection: Soil moisture effects on the atmospheric boundary layer in fully-coupled simulations". *Advances in Water Resources*, 30 (12), 2447–2466.

Maxwell, R. M. and L. E. Condon (2016). "Connections between groundwater flow and transpiration partitioning". *Science*, 353 (6297), 377–380.

Maxwell, R. M. and S. J. Kollet (2008). "Interdependence of groundwater dynamics and land-energy feedbacks under climate change". *Nature Geoscience*, 1 (10), 665–669.

Maxwell, R. M., J. K. Lundquist, J. D. Mirocha, S. G. Smith, C. S. Woodward, and A. F. B. Tompson (2011). "Development of a Coupled Groundwater-Atmosphere Model". *Monthly Weather Review*, 139 (1), 96–116.

McCabe, M. F., M. Rodell, D. E. Alsdorf, D. G. Miralles, R. Uijlenhoet, W. Wagner, A. Lucieer, R. Houborg, N. Verhoest, T. E. Franz, J. Shi, H. Gao, and E. F. Wood (2017). "The future of Earth observation in hydrology". *Hydrology and Earth System Sciences*, 21 (7), 3879–3914.

McDermid, S. S., L. O. Mearns, and A. C. Ruane (2017). "Representing agriculture in Earth System Models: Approaches and priorities for development". *Journal of Advances in Modeling Earth Systems*, 9 (5), 2230–2265.

Mekonnen, M. M. and A. Y. Hoekstra (2016). "Four billion people facing severe water scarcity". *Science Advances*, 2 (2), e1500323.

Miguez-Macho, G. and Y. Fan (2012). "The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration". *Journal of Geophysical Research*, 117 (D15), D15114.

Miguez-Macho, G., G. L. Stenchikov, and A. Robock (2004). "Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations". *Journal of Geophysical Research: Atmospheres*, 109 (D13), D13104.

Milly, P. C. D. and K. A. Dunne (2016). "Potential evapotranspiration and continental drying". *Nature Climate Change*, 6 (10), 946–949.

Miralles, D., T. Holmes, R. A. M. de Jeu, J. Gash, A. Meesters, and A. Dolman (2011). "Global land-surface evaporation estimated from satellite-based observations". *Hydrology and Earth System Sciences*, 15 (2), 453–469.

Miralles, D. G., M. J. van den Berg, A. J. Teuling, and R. A. M. de Jeu (2012). "Soil moisture-temperature coupling: A multiscale observational analysis". *Geophysical Research Letters*, 39 (21), L21707.

Miralles, D. G., A. J. Teuling, C. C. van Heerwaarden, and J. Vilà-Guerau de Arellano (2014). "Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation". *Nature Geoscience*, 7 (5), 345–349.

Miralles, D. G., P. Gentine, S. I. Seneviratne, and A. J. Teuling (2018). "Land – atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges". *Annals of the New York Academy of Sciences*, 1–17.

Moore, N. and S. Rojstaczer (2001). "Irrigation-Induced Rainfall and the Great Plains". *Journal of Applied Meteorology*, 40, 1297–1309.

Mueller, B., S. I. Seneviratne, C. Jimenez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. Dirmeyer, J. B. Fisher, Z. Guo, M. Jung, F. Maignan, M. F. McCabe, R. Reichle, M. Reichstein, M. Rodell, J. Sheffield, A. J. Teuling, K. Wang, E. F. Wood, and Y. Zhang (2011). "Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations". *Geophysical Research Letters*, 38 (6), 1–7.

Mueller, B., M. Hirschi, C. Jimenez, P. Ciais, P. A. Dirmeyer, A. J. Dolman, J. B. Fisher, M. Jung, F. Ludwig, F. Maignan, D. G. Miralles, M. F. McCabe, M. Reichstein, J. Sheffield, K. Wang, E. F. Wood, Y. Zhang, and S. I. Seneviratne (2013). "Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis". *Hydrology and Earth System Sciences*, 17 (10), 3707–3720.

Mueller, B. and S. I. Seneviratne (2012). "Hot days induced by precipitation deficits at the global scale." *Proceedings of the National Academy of Sciences of the United States of America*, 109 (31), 12398–12403.

— (2014). "Systematic land climate and evapotranspiration biases in CMIP5 simulations". *Geophysical Research Letters*, 41 (1), 128–134.

Müller-Schmied, H., S. Eisner, D. Franz, M. Wattenbach, F. T. Portmann, M. Flörke, and P. Döll (2014). "Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration". *Hydrology and Earth System Sciences*, 18, 3511–3538.

Müller-Schmied, H., L. Adam, S. Eisner, G. Fink, M. Flörke, H. Kim, T. Oki, F. T. Portmann, R. Reinecke, C. Riedel, Q. Song, J. Zhang, and P. Döll (2016). "Variations of global and continental water balance components as impacted by climate forcing uncertainty and human water use". *Hydrology and Earth System Sciences*, 20 (7), 2877–2898.

Naz, B. S., W. Kurtz, C. Montzka, W. Sharples, K. Goergen, J. Keune, H. Gao, A. Springer, H. Hendricks-Franssen, and S. J. Kollet (2018). "Improving soil moisture and runoff simulations over Europe using a high-resolution data-assimilation modeling framework". *Hydrology and Earth System Sciences Discussions*. DOI: `10.5194/hess-2018-24`.

Nazemi, A. and H. S. Wheater (2015). "On inclusion of water resource management in Earth system models; Part 1: Problem definition and representation of water demand". *Hydrology and Earth System Sciences*, 19 (1), 33–61.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007). "Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data". *Journal of Geophysical Research*, 112 (D7), 1–14.

Oki, T. and S. Kanae (2006). "Global Hydrological Cycles and World Water Resources". *Science*, 313 (5790), 1068–1072.

Oleson, K. W., G. Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, R. Stöckli, R. E. Dickinson, G. B. Bonan, S. Levis, A. Dai, and T. Qian (2008). "Improvements to the community land model and their impact on the hydrological cycle". *Journal of Geophysical Research: Biogeosciences*, 113 (1), G01021.

Oleson, K., Y. Dai, G. Bonan, M. Bosilovichm, R. Dickinson, P. Dirmeyer, F. Hoffman, P. Houser, S. Levis, G.-Y. Niu, P. Thornton, M. Vertenstein, Z.-L. Yang, and X. Zeng (2004). *Technical description of the community land model (CLM)*. Tech. Note TN-461+STR, 174. National Centre for Atmospheric Research.

Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994). "Verification, validation, and confirmation of numerical models in the earth sciences". *Science*, 263 (5147), 641–646.

Orlanski, I. (1975). "A rational subdivision of scales for atmospheric processes". *Bulletin of the American Meteorological Society*, 527–530.

Orth, R., E. Dutra, and F. Pappenberger (2016). "Improving weather predictability by including land surface model parameter uncertainty". *Monthly Weather Review*, 144 (4), 1551–1569.

Pal, J. S. and E. A. B. Eltahir (2001). "Pathways relating soil moisture conditions to future summer rainfall within a model of the land–atmosphere system". *Journal of Climate*, 14 (6), 1227–1242.

Pei, L., N. Moore, S. Zhong, A. D. Kendall, Z. Gao, and D. W. Hyndman (2016). "Effects of irrigation on summer precipitation over the United States". *Journal of Climate*, 29 (10), 3541–3558.

Perkins, S., L. Alexander, and J. Nairn (2012). "Increasing frequency, intensity and duration of observed global heatwaves and warm spells". *Geophysical Research Letters*, 39 (20).

Pokhrel, Y. N., S. Koirala, P. J.-F. Yeh, N. Hanasaki, L. Longuevergne, S. Kanae, and T. Oki (2015). "Incorporation of groundwater pumping in a global land surface model with the representation of human impacts". *Water Resources Research*, 51 (1), 78–96.

Prein, A., A. Gobiet, M. Suklitsch, H. Truhetz, N. Awan, K. Keuler, and G. Georgievski (2013). "Added value of convection permitting seasonal simulations". *Climate Dynamics*, 41 (9-10), 2655–2677.

Prein, A. F., W. Langhans, G. Fosser, A. Ferrone, N. Ban, K. Goergen, M. Keller, M. Tölle, O. Gutjahr, F. Feser, E. Brisson, S. J. Kollet, J. Schmidli, N. P. M. van Lipzig, and R. Leung (2015). "A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges". *Reviews of Geophysics*, 53 (2), 323–361.

Puma, M. J. and B. I. Cook (2010). "Effects of irrigation on global climate during the 20th century". *Journal of Geophysical Research: Atmospheres*, 115 (D16), D16120.

Quinn, P., K. Beven, and A. Culf (1995). "The introduction of macroscale hydrological complexity into land surface-atmosphere transfer models and the effect on planetary boundary layer development". *Journal of Hydrology*, 166 (3-4), 421–444.

Rahman, M. A.S. M., M. Sulis, and S. J. Kollet (2014). "The concept of dual-boundary forcing in land surface-subsurface interactions of the terrestrial hydrologic and energy cycles". *Water Resources Research*, 50 (11), 8531–8548.

Rahman, M. A.S. M., M. Sulis, and S. J. Kollet (2015). "The subsurface – land surface – atmosphere connection under convective conditions". *Advances in Water Resources*, 83, 240–249.

Rakovec, O., R. Kumar, J. Mai, M. Cuntz, S. Thober, M. Zink, S. Attinger, D. Schäfer, M. Schrön, and L. Samaniego (2016). "Multiscale and multivariate evaluation of water fluxes and states over European river basins". *Journal of Hydrometeorology*, 17 (1), 287–307.

Rast, M., J. Johannessen, and W. Mauser (2014). "Review of understanding of Earth's hydrological cycle: Observations, theory and modelling". *Surveys in Geophysics*, 35 (3), 491–513.

Ritter, B. and J.-F. Geleyn (1992). "A Comprehensive Radiation Scheme for Numerical Weather Prediction Models with Potential Applications in Climate Simulations". *Monthly Weather Review*, 120 (2), 303–325.

Rockel, B. (2015). "The Regional Downscaling Approach: a Brief History and Recent Advances". *Current Climate Change Reports*, 1 (1), 22–29.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009). "Satellite-based estimates of groundwater depletion in India". *Nature*, 460 (7258), 999–1002.

Rojas, R., L. Feyen, A. Dosio, and D. Bavera (2011). "Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations." *Hydrology and Earth System Sciences*, 15 (8), 2599–2620.

Rudolf, B., C. Beck, J. Grieser, and U. Schneider (2005). *Global precipitation analysis products of the GPCC*. Tech. Report. Global Precipitation Climatology Centre (GPCC) - Deutscher Wetterdienst, Offenbach am Main.

Rummukainen, M. (2016). "Added value in regional climate modeling". *Wiley Interdisciplinary Reviews: Climate Change*, 7 (1), 145–159.

Sacks, W. J., B. I. Cook, N. Buenning, S. Levis, and J. H. Helkowski (2009). "Effects of global irrigation on the near-surface climate". *Climate Dynamics*, 33 (2-3), 159–175.

Savenije, H. H. (1993). "Predictive model for salt intrusion in estuaries". *Journal of Hydrology*, 148 (1-4), 203–218.

Scanlon, B. R., C. C. Faunt, L. Longuevergne, R. C. Reedy, W. M. Alley, V. L. McGuire, and P. B. McMahon (2012). "Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley". *Proceedings of the National Academy of Sciences of the United States of America*, 109 (24), 9320–9325.

Scanlon, B. R., Z. Zhang, H. Save, A. Y. Sun, H. Müller-Schmied, L. P. H. van Beek, D. N. Wiese, Y. Wada, D. Long, R. C. Reedy, L. Longuevergne, P. Döll, and M. F. P. Bierkens (2018). "Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data". *Proceedings of the National Academy of Sciences of the United States of America*, 201704665. DOI: `10.1073/pnas.1704665115`.

Schaap, M. G. and F. J. Leij (1998). "Database-related accuracy and uncertainty of pedotransfer functions". *Soil Science*, 163, 765–779.

Schaller, M. F. and Y. Fan (2009). "River basins as groundwater exporters and importers: Implications for water cycle and climate modeling". *Journal of Geophysical Research*, 114 (D4), D04103.

Schär, C. and G. Jendritzky (2004). "Climate change: hot news from summer 2003." *Nature*, 432 (7017), 559–560.

Schättler, U. (2013). *A description of the nonhydrostatic regional COSMO model. Part V: Preprocessing: Initial and Boundary Data for the COSMO Model*. Tech. Report. Deutscher Wetterdienst, Offenbach am Main.

Schewe, J., J. Heinke, D. Gerten, I. Haddeland, N. W. Arnell, D. B. Clark, R. Dankers, S. Eisner, B. M. Fekete, F. J. Colón-González, S. N. Gosling, H. Kim, X. Liu, Y. Masaki, F. T. Portmann, Y. Satoh, T. Stacke, Q. Tang, Y. Wada, D. Wisser, T. Albrecht, K. Frieler, F. Piontek, L. Warszawski, and P. Kabat (2014). "Multimodel assessment of water scarcity under climate change." *Proceedings of the National Academy of Sciences of the United States of America*, 111 (9), 3245–3250.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Ziese (2011). *GPCC full data reanalysis version 6.0 at 0.5: Monthly land-surface precipitation from rain-gauges built on GTS-based and historic data*. Tech. Report. Global Precipitation Climatology Centre (GPCC) - Deutscher Wetterdienst, Offenbach am Main.

Seneviratne, S. I., D. Lüthi, M. Litschi, and C. Schär (2006). "Land–atmosphere coupling and climate change in Europe". *Nature*, 443 (7108), 205.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010). "Investigating soil moisture–climate interactions in a changing climate: A review". *Earth-Science Reviews*, 99 (3-4), 125–161.

Seuffert, G., P. Gross, C. Simmer, and E. F. Wood (2002). "The Influence of Hydrologic Modeling on the Predicted Local Weather: Two-Way Coupling of a Mesoscale Weather Prediction Model and a Land Surface Hydrologic Model". *Journal of Hydrometeorology*, 3 (5), 505–523.

Shrestha, P., M. Sulis, M. Masbou, S. J. Kollet, and C. Simmer (2014). "A scale-consistent Terrestrial Systems Modeling Platform based on COSMO, CLM and ParFlow." *Monthly Weather Review*, 9 (142), 3466–3483.

Shrestha, P., M. Sulis, C. Simmer, and S. J. Kollet (2015). "Impacts of grid resolution on surface energy fluxes simulated with an integrated surface-groundwater flow model". *Hydrology and Earth System Sciences*, 19 (10), 4317–4326.

Siebert, S. and P. Döll (2010). "Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation". *Journal of Hydrology*, 384 (3-4), 198–217.

Siebert, S., J. Burke, J.-M. Faures, K. Frenken, J. Hoogeveen, P. Döll, and F. T. Portmann (2010). "Groundwater use for irrigation–a global inventory". *Hydrology and Earth System Sciences*, 14 (10), 1863–1880.

Simmer, C., I. Thiele-Eich, M. Masbou, W. Amelung, H. Bogena, S. Crewell, B. Diekkrüger, F. Ewert, H.-J. Hendricks-Franssen, J. A. Huisman, A. Kemna, N. Klitzsch, S. J. Kollet, M. Langensiepen, U. Löhnert, M. A.S. M. Rahman, U. Rascher, K. Schneider, J. Schween, Y. Shao, P. Shrestha, M. Stiebler, M. Sulis, J. Vanderborght, H. Vereecken, J. van der Kruk, G. Waldhoff, and T. Zerenner (2015). "Monitoring and modeling the terrestrial system from pores to catchments: the transregional collaborative research center on patterns in the soil–vegetation–atmosphere system". *Bulletin of the American Meteorological Society*, 96 (10), 1765–1787.

Soden, B. J. and I. M. Held (2006). "An assessment of climate feedbacks in coupled ocean–atmosphere models". *Journal of Climate*, 19 (14), 3354–3360.

Sörensson, A. A. and R. C. Ruscica (2018). "Intercomparison and Uncertainty Assessment of Nine Evapotranspiration Estimates Over South America". *Water Resources Research*, 54, 2891–2908.

Soylu, M. E., E. Istanbulluoglu, J. D. Lenters, and T. Wang (2011). "Quantifying the impact of groundwater depth on evapotranspiration in a semi-arid grassland region". *Hydrology and Earth System Sciences*, 15 (3), 787–806.

Staniforth, A. (1997). "Regional modeling: A theoretical discussion". *Meteorology and Atmospheric Physics*, 63 (1-2), 15–29.

Stéfanon, M., P. Drobinski, F. D'Andrea, C. Lebeaupin-Brossier, and S. Bastin (2013). "Soil moisture-temperature feedbacks at meso-scale during summer heat waves over Western Europe". *Climate Dynamics*, 42 (5-6), 1309–1324.

Sulis, M., C. Paniconi, and M. Camporese (2011). "Impact of grid resolution on the integrated and distributed response of a coupled surface-subsurface hydrological model for the des Anglais catchment, Quebec". *Hydrological Processes*, 25 (12), 1853–1865.

Sulis, M., J. L. Williams, P. Shrestha, M. Diederich, C. Simmer, S. J. Kollet, and R. M. Maxwell (2017). "Coupling groundwater, vegetation, and atmospheric processes: a comparison of two integrated models". *Journal of Hydrometeorology*, 18 (5), 1489–1511.

Sulis, M., J. Keune, P. Shrestha, C. Simmer, and S. J. Kollet (2018). "Quantifying the impact of subsurface-land surface physical processes on the predictive skill of subseasonal mesoscale atmospheric simulations". *Journal of Geophysical Research: Atmospheres*, 123. DOI: 10.1029/2017JD028187.

Swenson, S. and D. Lawrence (2014). "Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data". *Journal of Geophysical Research: Atmospheres*, 119 (17).

Taylor, C. M., R. A. M. de Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo (2012). "Afternoon rain more likely over drier soils". *Nature*, 489 (7416), 423.

Taylor, R. G., B. R. Scanlon, P. Döll, M. Rodell, L. P. H. van Beek, Y. Wada, L. Longuevergne, M. Leblanc, J. S. Famiglietti, M. Edmunds, M. Konikow, T. R. Green, J. Chen, M. Taniguchi, M. F. P. Bierkens, A. MacDonald, Y. Fan, R. M. Maxwell, Y. Yechieli, J. J. Gurdak, D. M. Allen, M. Shamsudduha, K. Hiscock, P. J.-F. Yeh, I. Holman, and H. Treidel (2013). "Ground water and climate change". *Nature Climate Change*, 3 (4), 322.

Thiery, W., E. L. Davin, D. M. Lawrence, A. L. Hirsch, M. Hauser, and S. I. Seneviratne (2017). "Present-day irrigation mitigates heat extremes". *Journal of Geophysical Research: Atmospheres*, 122 (3), 1403–1422.

Tiedtke, M. (1989). "A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models". *Monthly Weather Review*, 117 (8), 1779–1800.

Trambauer, P., E. Dutra, S. Maskey, M. Werner, F. Pappenberger, L. P. H. van Beek, and S. Uhlenbrook (2014). "Comparison of different evaporation estimates over the African continent". *Hydrology and Earth System Sciences*, 18 (1), 193–212.

Tuinenburg, O. A. and J. P. R. de Vries (2017). "Irrigation Patterns Resemble ERA-Interim Reanalysis Soil Moisture Additions". *Geophysical Research Letters*, 44 (20).

Turner, B. L., E. F. Lambin, and A. Reenberg (2007). "The emergence of land change science for global environmental change and sustainability". *Proceedings of the National Academy of Sciences of the United States of America*, 104 (52), 20666–20671.

Valcke, S. (2013). "The OASIS3 coupler: a European climate modelling community software". *Geoscientific Model Development*, 6 (2), 373–388.

van der Velde, M., G. Wriedt, and F. Bouraoui (2010). "Estimating irrigation use and effects on maize yield during the 2003 heatwave in France". *Agriculture, Ecosystems & Environment*, 135 (1-2), 90–97.

Van Genuchten, M. T. (1980). "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils". *Soil Science Society of America Journal*, 44 (5), 892–898.

Vautard, R., P. Yiou, F. D'Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, and Y. Fan (2007). "Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit". *Geophysical Research Letters*, 34 (7), L07711.

Vautard, R., A. Gobiet, D. Jacob, M. Belda, A. Colette, M. Déqué, J. Fernández, M. García-Díez, K. Goergen, I. Güttler, T. Halenka, T. Karacostas, E. Katragkou, K. Keuler, S. Kotlarski, S. Mayer, E. van Meijgaard, G. Nikulin, M. Patarčić, J. Scinocca, S. Sobolowski, M. Suklitsch, C. Teichmann, K. Warrach-Sagi, V. Wulfmeyer, and P. Yiou (2013). "The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project". *Climate Dynamics*, 41 (9-10), 2555–2575.

Veldkamp, T. I. E., F. Zhao, P. J. Ward, H. de Moel, J. C.J. H. Aerts, H. Müller-Schmied, F. T. Portmann, Y. Masaki, Y. Pokhrel, X. Liu, Y. Satoh, D. Gerten, S. N. Gosling, J. Zaherpour, and Y. Wada (2018). "Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study". *Environmental Research Letters*, 13 (5), 055008.

Voisin, N., H. Li, D. Ward, M. Huang, M. Wigmosta, and L. R. Leung (2013). "On an improved sub-regional water resources management representation for integration into earth system models". *Hydrology and Earth System Sciences*, 17 (9), 3605–3622.

von Storch, H., H. Langenberg, and F. Feser (2000). "A Spectral Nudging Technique for Dynamical Downscaling Purposes". *Monthly Weather Review*, 128 (10), 3664–3673.

von Storch, H. and F. W. Zwiers (2002). *Statistical analysis in climate research*. Citeseer.

Vörösmarty, C. J., P. Green, J. Salisbury, and R. B. Lammers (2000). "Global Water Resources: Vulnerability from Climate Change and Population Growth". 289 (5477), 284–288.

Vukicevic, T. and R. M. Errico (1990). "The influence of artificial and physical factors upon predictability estimates using a complex limited-area model". *Monthly Weather Review*, 118 (7), 1460–1482.

Wada, Y., D. Wisser, and M. F. P. Bierkens (2014). "Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources". *Earth System Dynamics Discussions*, 5 (1), 15–40.

Wada, Y., I. E. M. de Graaf, and L. P. H. van Beek (2016a). "High-resolution modeling of human and climate impacts on global water resources". *Journal of Advances in Modeling Earth Systems*, 8 (2), 735–763.

Wada, Y., L. P. H. van Beek, and M. F. P. Bierkens (2012). "Nonsustainable groundwater sustaining irrigation: A global assessment". *Water Resources Research*, 48, W00L06.

Wada, Y., D. Wisser, S. Eisner, M. Flörke, D. Gerten, I. Haddeland, N. Hanasaki, Y. Masaki, F. T. Portmann, T. Stacke, Z. Tessler, and J. Schewe (2013). "Multimodel projections and uncertainties of irrigation water demand under climate change". *Geophysical Research Letters*, 40 (17), 4626–4632.

Wada, Y., J. T. Reager, B. F. Chao, J. Wang, M.-H. Lo, C. Song, Y. Li, and A. S. Gardner (2016b). "Recent Changes in Land Water Storage and its Contribution to Sea Level Variations". *Surveys in Geophysics*, 1–22.

Wada, Y., M. F. P. Bierkens, A. de Roo, P. A. Dirmeyer, J. S. Famiglietti, N. Hanasaki, M. Konar, J. Liu, H. Müller-Schmied, T. Oki, Y. Pokhrel, M. Sivapalan, T. J. Troy, A. I.J. M. van Dijk, T. van Emmerik, M. H. J. Van Huijgevoort, H. A. J. Van Lanen,

C. J. Vörösmarty, N. Wanders, and H. Wheater (2017). "Human–water interface in hydrological modelling: current status and future directions". *Hydrology and Earth System Sciences*, 21 (8), 4169–4193.

Wagener, T., M. Sivapalan, P. A. Troch, B. L. McGlynn, C. J. Harman, H. V. Gupta, P. Kumar, P. Rao, N. B. Basu, and J. S. Wilson (2010). "The future of hydrology: An evolving science for a changing world". *Water Resources Research*, 46 (5), W05301.

Waldron, K. M., J. Paegle, and J. D. Horel (1996). "Sensitivity of a spectrally filtered and nudged limited-area model to outer model options". *Monthly Weather Review*, 124 (3), 529–547.

Walko, R. L., L. E. Band, J. Baron, T. G. F. Kittel, R. Lammers, T. J. Lee, D. Ojima, R. A. Pielke, C. M. Taylor, C. Tague, C. J. Tremback, and P. L. Vidale (2000). "Coupled Atmosphere–Biophysics–Hydrology Models for Environmental Modeling". *Journal of Applied Meteorology*, 39 (6), 931–944.

Wanders, N. and Y. Wada (2015). "Human and climate impacts on the 21st century hydrological drought". *Journal of Hydrology*, 526, 208–220.

Wang, A., X. Zeng, and D. Guo (2016). "Estimates of Global Surface Hydrology and Heat Fluxes from the Community Land Model (CLM4.5) with Four Atmospheric Forcing Datasets". *Journal of Hydrometeorology*, 17, 2493–2510.

Wang-Erlandsson, L., I. Fetzer, P. W. Keys, R. J. van der Ent, H. H. G. Savenije, and L. J. Gordon (2017). "Remote land use impacts on river flows through atmospheric teleconnections". *Hydrology and Earth System Sciences Discussions*. DOI: 10.5194/hess-2017-494.

Warner, J. C., N. Perlin, and E. D. Skyllingstad (2008). "Using the Model Coupling Toolkit to couple earth system models". *Environmental Modelling & Software*, 23 (10-11), 1240–1249.

Warner, T. T., R. A. Peterson, and R. E. Treadon (1997). "A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction". *Bulletin of the American Meteorological Society*, 78 (11), 2599–2618.

Wei, J., H. Su, and Z.-L. Yang (2016). "Impact of moisture flux convergence and soil moisture on precipitation: a case study for the southern United States with implications for the globe". *Climate Dynamics*, 46 (1-2), 467–481.

Weisheimer, A., F. J. Doblas-Reyes, T. Jung, and T. N. Palmer (2011). "On the predictability of the extreme summer 2003 over Europe". *Geophysical Research Letters*, 38 (5), L05704.

Wicker, L. J. and W. C. Skamarock (2002). "Time-splitting methods for elastic models using forward time schemes". *Monthly Weather Review*, 130 (8), 2088–2097.

Willmott, C. J. and K. Matsuura (2001). *Terrestrial air temperature and precipitation: monthly and annual climatologies (version 3.02)*. URL: http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_clim2.html (visited on 08/04/2018).

Willmott, C. J. and S. M. Robeson (1995). "Climatologically aided interpolation (CAI) of terrestrial air temperature". *International Journal of Climatology*, 15 (2), 221–229.

Wilson, K., A. Goldstein, E. Falge, M. Aubinet, D. Baldocchi, P. Berbigier, C. Bernhofer, R. Ceulemans, H. Dolman, C. Field, A. Grelle, A. Ibrom, B. E. Law, A. Kowalski,

T. Meyers, J. Moncrieff, R. Monson, W. Oechel, J. Tenhunen, R. Valentini, and S. Verma (2002). "Energy balance closure at FLUXNET sites". *Agricultural and Forest Meteorology*, 113 (1), 223–243.

Wood, E. F., J. K. Roundy, T. J. Troy, L. P. H. van Beek, M. F. P. Bierkens, E. Blyth, A. de Roo, P. Döll, M. Ek, J. S. Famiglietti, D. Gochis, N. van de Giesen, P. Houser, P. R. Jaffeé, S. J. Kollet, B. Lehner, D. P. Lettenmaier, C. Peters-Lidard, M. Sivapalan, J. Sheffield, A. Wade, and P. Whitehead (2011). "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water". *Water Resources Research*, 47 (5), W05301.

Wouters, H., K. De Ridder, L. Poelmans, P. Willems, J. Brouwers, P. Hosseinzadehtalaei, H. Tabari, S. Vanden Broucke, N. P. M. van Lipzig, and M. Demuzere (2017). "Heat stress increase under climate change twice as large in cities as in rural areas: A study for a densely populated midlatitude maritime region". *Geophysical Research Letters*, 44 (17), 8997–9007.

Yeh, P. J.-F. and E. A. B. Eltahir (2005). "Representation of water table dynamics in a land surface scheme. Part I: Model development". *Journal of Climate*, 18 (12), 1861–1880.

York, J. P., M. Person, W. J. Gutowski, and T. C. Winter (2002). "Putting aquifers into atmospheric simulation models: an example from the Mill Creek Watershed, northeastern Kansas". *Advances in Water Resources*, 25 (2), 221–238.

Zaitchik, B. F., M. Rodell, and R. H. Reichle (2008). "Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin". *Journal of Hydrometeorology*, 9 (3), 535–548.

Zaitchik, B. F., A. K. Macalady, L. R. Bonneau, and R. B. Smith (2006). "Europe's 2003 heat wave: a satellite view of impacts and land–atmosphere feedbacks". *International Journal of Climatology*, 26 (6), 743–769.

Zhang, Y., J. L. Peña-Arancibia, T. R. McVicar, F. H. Chiew, J. Vaze, C. Liu, X. Lu, H. Zheng, Y. Wang, Y. Y. Liu, D. G. Miralles, and M. Pan (2016). "Multi-decadal trends in global terrestrial evapotranspiration and its components". *Scientific Reports*, 6, 19124.

BONNER METEOROLOGISCHE ABHANDLUNGEN

Herausgegeben vom Meteorologischen Institut der Universität Bonn durch Prof. Dr. H. FLOHN (Hefte 1-25), Prof. Dr. M. HANTEL (Hefte 26-35), Prof. Dr. H.-D. SCHILLING (Hefte 36-39), Prof. Dr. H. KRAUS (Hefte 40-49), ab Heft 50 durch Prof. Dr. A. HENSE.

Heft 1-63: siehe `http://www.meteo.uni-bonn.de/bibliothek/bma`

64-77: open access, verfügbar unter `https://uni-bn.de/kpSDaQffe1`

Heft 64: **Michael Weniger**: Stochastic parameterization: a rigorous approach to stochastic three-dimensional primitive equations, 2014, 148 S. + XV.

Heft 65: **Andreas Röpnack**: Bayesian model verification: predictability of convective conditions based on EPS forecasts and observations, 2014, 152 S. + VI.

Heft 66: **Thorsten Simon**: Statistical and Dynamical Downscaling of Numerical Climate Simulations: Enhancement and Evaluation for East Asia, 2014, 48 S. + VII. + Anhänge

Heft 67: **Elham Rahmani**: The Effect of Climate Change on Wheat in Iran, 2014, [erschienen] 2015, 96 S. + XIII.

Heft 68: **Pablo A. Saavedra Garfias**: Retrieval of Cloud and Rainwater from Ground-Based Passive Microwave Observations with the Multi-frequency Dual-polarized Radiometer ADMIRARI, 2014, [erschienen] 2015, 168 S. + XIII.

Heft 69: **Christoph Bollmeyer**: A high-resolution regional reanalysis for Europe and Germany - Creation and Verification with a special focus on the moisture budget, 2015, 103 S. + IX.

Heft 70: **A S M Mostaquimur Rahman**: Influence of subsurface hydrodynamics on the lower atmosphere at the catchment scale, 2015, 98 S. + XVI.

Heft 71: **Sabrina Wahl**: Uncertainty in mesoscale numerical weather prediction: probabilistic forecasting of precipitation, 2015, 108 S.

Heft 72: **Markus Übel**: Simulation of mesoscale patterns and diurnal variations of atmospheric CO_2 mixing ratios with the model system TerrSysMP-CO_2, 2015, [erschienen] 2016, 158 S. + II

Heft 73: **Christian Bernardus Maria Weijenborg**: Characteristics of Potential Vorticity anomalies associated with mesoscale extremes in the extratropical troposphere, 2015, [erschienen] 2016, 151 S. + XI

Heft 74: ***Muhammad Kaleem***: A sensitivity study of decadal climate prediction to aerosol variability using ECHAM6-HAM (GCM), 2016, 98 S. + XII

Heft 75: ***Theresa Bick***: 3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale, 2016, [erschienen] 2017, 96 S. + IX

Heft 76: ***Zied Ben Bouallegue***: Verification and post-processing of ensemble weather forecasts for renewable energy applications, 2017, 119 S.

Heft 77: ***Julia Lutz***: Improvements and application of the STatistical Analogue Resampling Scheme STARS, 2016, [erschienen] 2017, 103 S.

Heft 78: ***Benno Michael Thoma***: Palaeoclimate Reconstruction in the Levant and on the Balkans, 2016, [erschienen] 2017, XVI, 266 S.

Heft 79: ***Ieda Pscheidt***: Generating high resolution precipitation conditional on rainfall observations and satellite data, 2017, V, 173 S.

Heft 80: ***Tanja Zerenner***: Atmospheric downscaling using multi-objective genetic programming, 2016, [erschienen] 2017, X, 191 S.

Heft 81: ***Sophie Stolzenberger***: On the probabilistic evaluation of decadal and paleoclimate model predictions, 2017, IV, 122 S.

Heft 82: ***Insa Thiele-Eich***: Flooding in Dhaka, Bangladesh, and the challenge of climate change, 2017, V, 158 S.

Heft 83: ***Liselotte Bach***: Towards a probabilistic regional reanalysis for Europe, 2017 [erschienen] 2018, VI, 114 S.

Heft 84: ***Yen-Sen Lu***: Propagation of land surface model uncertainties in terrestrial system states, 2017, [erschienen] 2018, X, 120 S.

Heft 85: ***Rüdiger Hewer***: Stochastic physical models for wind fields and precipitation extremes, 2018, 99 S.

Heft 86: ***Sebastian Knist***: Land-atmosphere interactions in multiscale regional climate change simulations over Europe, 2018, 147 S.

Heft 87: ***Jessica Keune***: Integrated terrestrial simulations at the continental scale: Impact of groundwater dynamics and human water use on groundwater-to-atmosphere feedbacks during the European heatwave in 2003, 2019, IX, 172 S.