High-resolution fully-coupled atmospheric–hydrological modeling: a cross-compartment regional water and energy cycle evaluation

Benjamin Fersch[1], Alfonso Senatore[2], Bianca Adler[3], Joël Arnault[1], Matthias Mauder[1], Katrin Schneider[1], Ingo Völksch[1], and Harald Kunstmann[1]

[1]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research (IMK-IFU), Garmisch-Partenkirchen, Germany
[2]University of Calabria, Department of Environmental and Chemical Engineering, Rende (CS), Italy
[3]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research (IMK-TRO), Karlsruhe, Germany

Correspondence to: Benjamin Fersch (fersch@kit.edu)

Abstract. The land surface and the atmospheric boundary layer are closely intertwined with respect to the exchange of water, trace gases and energy. Nonlinear feedback and scale dependent mechanisms are obvious by observations and theories. Modeling instead is often narrowed to single compartments of the terrestrial system or bound to traditional viewpoints of definite scientific disciplines. Coupled terrestrial hydrometeorological modeling systems attempt to overcome these limitations to achieve a better integration of the processes relevant for regional climate studies and local area weather prediction. This study examines the ability of the hydrologically enhanced version of the Weather Research and Forecasting Model (WRF-Hydro) to reproduce the regional water cycle by means of a two-way coupled approach and assesses the impact of hydrological coupling with respect to a traditional regional atmospheric model setting. It includes the observation-based calibration of the hydrological model component (offline WRF-Hydro) and a comparison of the classic WRF and the fully coupled WRF-Hydro models both with identical calibrated parameter settings for the land surface model (Noah-MP). The simulations are evaluated based on extensive observations at the Pre-Alpine Terrestrial Environmental Observatory (TERENO Pre-Alpine) for the Ammer (600 km^2) and Rott (55 km^2) river catchments in southern Germany, covering a five month period (Jun–Oct 2016). The sensitivity of 7 land surface parameters is tested using the *Latin-Hypercube One-factor-At-a-Time* (LH-OAT) method and 6 sensitive parameters are subsequently optimized for 6 different subcatchments, using the Model-Independent *Parameter Estimation and Uncertainty Analysis software* (PEST). The calibration of the offline WRF-Hydro gives Nash-Sutcliffe efficiencies between 0.56 and 0.64 and volumetric efficiencies between 0.46 and 0.81 for the six subcatchments. The comparison of classic WRF and fully coupled WRF-Hydro, both using the calibrated parameters from the offline model, shows only tiny alterations for radiation and precipitation but considerable changes for moisture- and heat fluxes. By comparison with TERENO Pre-Alpine observations, the fully coupled model slightly outperforms the classic WRF with respect to evapotranspiration, sensible and ground heat flux, near surface mixing ratio, temperature, and boundary layer profiles of air temperature. The subcatchment-based water budgets show uniformly directed variations for evapotranspiration, infiltration excess and percolation whereas soil moisture and precipitation change randomly.

1 Introduction

The intertwined exchange of water and energy fluxes at the land–atmosphere interface determines hydrological processes on a multitude of spatial and temporal scales. Its appropriate formulation and implementation into model systems is a prerequisite for climate- and land use change impact investigations. Both, terrestrial and atmospheric processes need to be considered. Fully coupled hydrological–atmospheric model system have recently been developed and comprise the most relevant Earth system components. Comprehensive and concerted evaluation of these coupled modeling systems is required to assess the current limits and potential in Earth system science. This study accordingly focuses on the evaluation of a fully-coupled atmospheric–hydrological model across the various compartments of the water and energy cycle.

As shown by Ning et al. (2019), in scientific literature, the topic of coupled hydrological–atmospheric modeling is constantly gaining interest. Several physically based, fully coupled hydrological–atmospheric models have been developed by the scientific community over the past 15 years, addressing non-linear cross-compartment feedback and fostering a closed representation of regional water and energy cycles (e.g., Shrestha et al., 2014; Butts et al., 2014; Gochis et al., 2016; Soltani et al., 2019). Comprehensive reviews on the history of fully coupled hydrometeorological models and their application can be found in Wagner et al. (2016), Senatore et al. (2015), and Ning et al. (2019). Typically, these models are amalgamations of preexisting subject-specific algorithms of varying complexity, with land surface models being the common thread. Recent applications of fully coupled models show promising results in improving spatial pattern dynamics and area integrals of regional water budgets. However, the research field is far away from maturity and many further studies are required.

Using ParFlow coupled with the Community Land Model, Maxwell and Kollet (2008) found that for the U.S. Oklahoma southern Great Planes, groundwater depth governs the sensitivity of regions to variations in temperature and precipitation. Larsen et al. (2016a) reported that by accounting for shallow groundwater in the fully coupled model MIKE SHE, summer evapotranspiration results improved for a study over Kansas, USA. While several studies highlight the importance of lateral hydrological processes for the improved simulation of soil moisture (e.g., Wagner et al., 2016; Larsen et al., 2016a), the sensitivity for land subsurface–surface–planetary boundary layer feedback and precipitation generation is less pronounced, especially for the humid regions with strong synoptic forcing (e.g., Butts et al., 2014; Barlage et al., 2015; Arnault et al., 2018; Rummler et al., 2018; Sulis et al., 2018).

Coupled modeling studies often focus on single objective variables for validation (like, e.g., discharge, evapotranspiration or soil moisture) or restrict their analysis to describing only the changes in simulation results without any comparison to observations. Targeting single variables can result in the problem of equifinality where model realizations are proven skillful for a single aspect, yet possibly being wrong for several others. To investigate if a certain model or model configuration can provide improved realism, the limited perspective of single or few variable evaluations needs to be abandoned (García-Díez et al., 2015). To overcome the dilemma, fully coupled simulations should be validated and evaluated with respect to as many independent observations as possible. However, the scales of simulations and observations need to match. For catchment-scale coupled hydrometeorological models, most of the global data products (e.g., from satellite) are rather coarse. Regional observatories with integrative measurements of the subsurface to boundary layer fluxes and states provide a sound basis for a

holistic evaluation. In the recent past, several efforts have been undertaken to create comprehensive observation sets that allow for subsurface to atmosphere integrated studies of water and energy fluxes for small to medium scale river catchments. The most prominent activities for Europe are HOAL (Hydrological Open Air Observatory, Blöschl et al., 2016), HOBE (The Danish Hydrological Observatory, Jensen and Refsgaard, 2018), LAFO (Land-atmosphere feedback observatory, Spath et al., 2018), and TERENO (TERestrial ENvironmental Observatories, Zacharias et al., 2011). Although two of them address hydrology in their names, land–atmosphere interaction is a central research item for all of these observatories.

Our study presents a concept to improve the physical realism of regional dynamical hydrometeorological simulations not only by taking into account lateral water redistribution processes on the land surface and their coupled feedback with the planetary boundary layer but also by evaluating the simulated water and energy budgets with comprehensive observations. We calibrate the land surface model that is used in the coupled modeling system based on discharge observations of several subcatchments and thus rely on a variable that integrates the hydrological behavior of the whole upstream area. In a classic local area modeling study, we could only tune land surface parameters based on station observations which would be less straightforward with respect to the different scales of simulation and observation. We investigate how well the hydrologically enhanced, fully-coupled model mimics observations for different compartments of the hydrological and the associated energy cycle. We evaluate the effect of bidirectional hydrological–atmospheric model coupling with respect to 1) the land surface energy flux partitioning and 2) to the different compartments of the hydrological cycle. To that end, we perform uncoupled and fully coupled simulations with the hydrologically enhanced version of the Weather Research and Forecasting modeling system (WRF-Hydro, Gochis et al., 2016) for the Ammer river catchment region, located in southern Bavaria, Germany. We utilize convection resolving resolution of 1 km^2 for the atmospheric part together with a 100 by 100 m hydrological subgrid. For validation, we employ a rich and comprehensive dataset consisting of observations from the TERENO Pre-Alpine observatory (Kiese et al., 2018), enhanced by data from the ScaleX field campaign (Wolf et al., 2016), complemented by further local providers.

2 Methods and data

2.1 Study area

The study area covers the two medium sized river catchments Ammer (600 km^2) and Rott (55 km^2), located in southern Bavaria, Germany (Fig. 1). The hydromorphic characteristics of this Alpine front-range region were formed during the last glacial and predominantly feature Gley- Cambi-, and Histosols on top of carbon based gravel deposits. Elevations range from above 2300 m ASL in the south, down to 533 m ASL at the outlet towards lake Ammer. Landcover is dominated by meadows and forests. The proportion of forests rises from about 20 % in the north to 57 % in the Alpine part of the catchment (Fetzer et al., 1986). Due to the climatic conditions, crops are only of minor importance and are only prevalent in the lower part. Mean annual precipitation exhibits a gradient from 950 mm close to lake Ammer to more than 2000 mm in the mountains. Mean annual evapotranspiration shows no distinct correlation with elevation and ranges from 300 mm at the sparsely vegetated mountain slopes to 500–600 mm

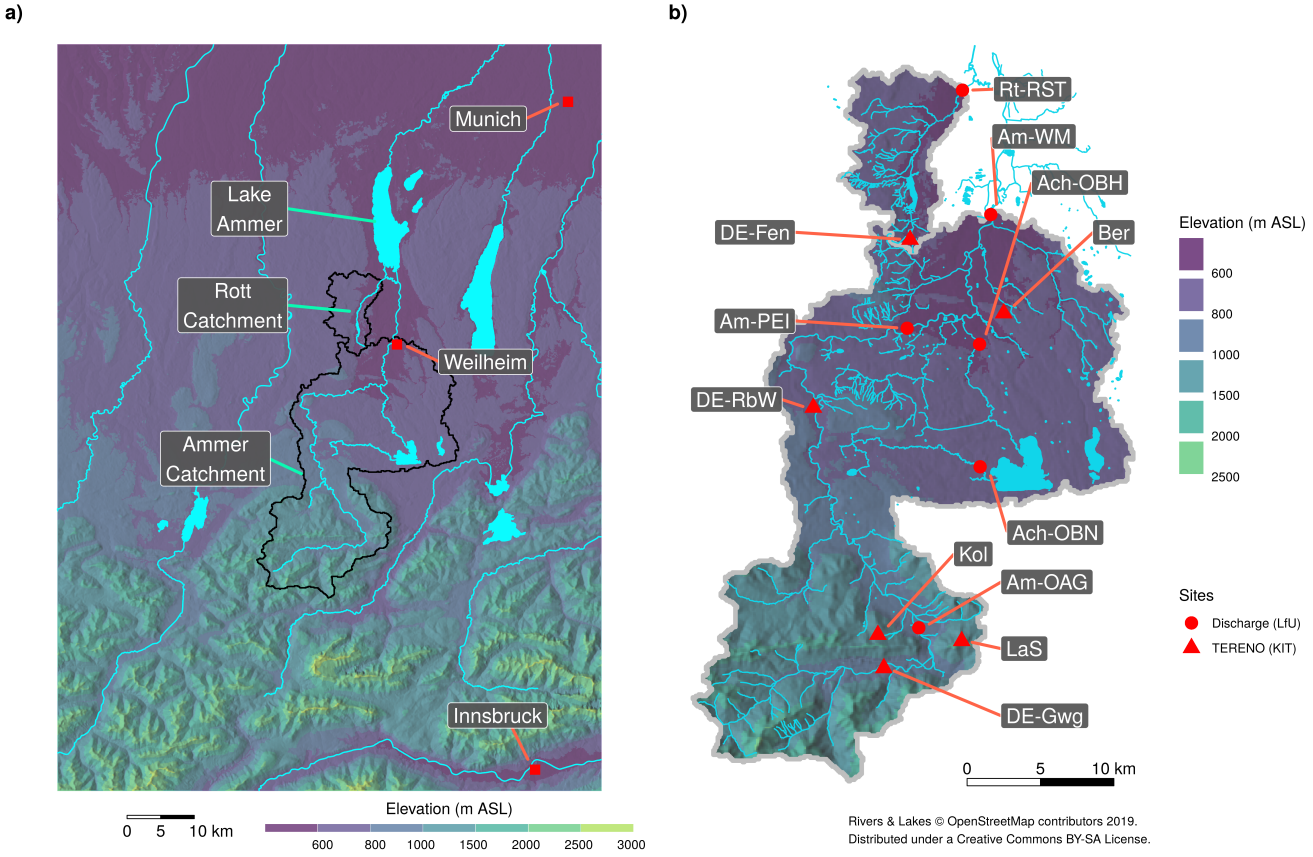

Figure 1. a) elevation map for model domain 3 and location of Ammer and Rott river catchments. b) catchment detail and observation locations. Abbreviations – river gauges maintained by the Bavarian Environmental Agency (LfU): Rt-RST = Rott-Raisting, Am-WM = Ammer-Weilheim, Am-PEI = Ammer-Peißenberg, Am-OAG = Ammer-Oberammergau, Ach-OBH = Ach-Oberhausen, Ach-OBN = Ach-Obernach; TERENO Pre-Alpine sites: DE-Fen = Peißenberg-Fendt, Ber = Oberhausen-Berg, DE-RbW = Rottenbuch, Kol = Oberammergau-Kolbensattel (mountain station), Las = Oberammergau-Laber (mountain station), DE-Gwg = Ettal-Graswang. Elevation is derived from ASTER GDEM.

for the rest of the catchment. Mean annual temperature ranges from about 7 to 4 °C between the lower and the upper parts of the catchment.

The study region is part of the German Terrestrial Environmental Observatory program (TERENO, Zacharias et al., 2011), an initiative for the long-term monitoring of climate environmental variables. Our study is bound to the multidisciplinary field experiment and observation campaign *ScaleX* (Wolf et al., 2016) that took place at the TERENO Pre-Alpine DE-Fen site (Fig. 1), in the summers of 2015 and 2016.

2.2 Observation data

The study region was selected to cover the Helmholtz Pre-Alpine TERrestrial ENvironmental Observatory (TERENO Pre-Alpine) located in the foothills of the Bavarian Alps of southern Germany. TERENO Pre-Alpine features observations for the range of compartments of the terrestrial hydrometeorological cycle. The observatory has been designed for long-term monitoring of climatological and ecological variables. A detailed description of the concept is available in Kiese et al. (2018). Figure 1b provides an overview of the measurement sites that comprise standard climatological, eddy-covariance, lysimeter, soil moisture, groundwater and discharge observations.

The observed sensible and latent heat fluxes presented in this study are determined by means of tower-based eddy-covariance measurements, which are operated on a long-term basis. These installations comprise a CSAT3 sonic anemometer (Campbell Scientific Inc., Logan, UT) at the three main sites (DE-Fen, DE-RbW, DE-Gwg) of TERENO Pre-Alpine and a LI-7500 infrared gas analyser at DE-Fen and DE-Gwg, while DE-RbW is equipped with a LI-7200 gas analyser (Licor Biosciences Inc., Lincoln, NE). The measurement height of these systems is 3.3 m above ground. High-frequency data from these instruments are recorded digitally on a Campbell CR3000 data logger. The fluxes are computed in the field every day with an industrial PC using the eddy-covariance software TK3 (Mauder and Foken, 2015), including corrections for misalignment of the anemometer using the double rotation methods (Wilczak et al., 2001), humidity influences on the sonic temperature measurement (Schotanus et al., 1983), spectral losses due to path averaging and sensor separation (Moore, 1986), and density fluctuations (Webb et al., 1980). Automated quality control and uncertainty assessment is applied in accordance with Mauder et al. (2013), which extends the test of Foken and Wichura (1996) by an additional spike test on the high-frequency data, a test on the interdependence of fluxes due to the flux corrections and a test on the representativeness of the flux footprint (Kormann and Meixner, 2001). Moreover, an energy-balance-closure-adjustment method, which is based on the daily energy balance ratio is applied to daytime sensible and latent heat flux data under the condition that the Bowen ratio is preserved (Mauder et al., 2013).

Lysimeter data is available for three of the TERENO Pre-Alpine sites (DE-Fen, DE-RbW, DE-Gwg). The measurements are separated for representative treatments of extensive and intensive grassland management, in accordance with the local farmer's cutting and fertilizer management (Fu et al., 2017). For this study, data derived from six control lysimeters per site are taken into account (i.e., lysimeters that were excavated at adjacent grassland sites nearby the experimental site). For each lysimeter, precipitation, evapotranspiration and groundwater recharge (percolation) is calculated from the variations in total weight and the changes in water volume of the corresponding water tank. Obvious outliers in the weight measurements are removed above thresholds of 1000 g min^{-1} and 200 g min^{-1} for the weight changes of the lysimeters and water tanks, respectively. Furthermore,

for separation of signal and noise, the *Adaptive Window and Adaptive Threshold filter* (AWAT, Peters et al., 2014) is applied to the time series of weight changes of each individual lysimeter and corresponding water tank at a temporal resolution of 1 min. The procedure applied in this study is further described by Fu et al. (2017).

A wireless sensor network at the DE-Fen site, consisting of 55 profiles (5, 20, 50 cm), provides soil moisture information for a grassland area of roughly 12 ha. The measurement devices are spade-shaped ring oscillator electromagnetic permittivity sensors (Truebner SMT 100, Bogena et al., 2017) with a vertical representativeness of about 3 cm. Additional information on sensor calibration and the conversion of permittivity into volumetric water content is available in Fersch et al. (2018).

Within the course of the ScaleX campaign (June-August 2016, Wolf et al., 2016), a scanning microwave radiometer (HATPRO, Humidity and temperature profiler, Rose et al., 2005) provided information on temperature and humidity profiles as well as integrated water vapor and liquid water path. The instrument measures sky brightness temperature at 14 frequencies, seven are distributed between 22.235 and 31.4 GHz along the wing of the 22.235 water vapor line and seven between 51.26 and 58 GHz along the wing of the 60 GHz oxygen absorption complex. Information on atmospheric variables is obtained from the measured brightness temperatures with a retrieval algorithm from the University of Cologne (Löhnert and Crewell, 2003; Löhnert et al., 2009). For the retrieval creation a set of around 14000 radiosonde profiles, measured at Munich (station at 489 m ASL) between 1990 and 2014, is used. While integrated water vapor has an accuracy of less than 1 kg m^{-2} (Pospichal and Crewell, 2007), the vertical resolution of humidity is low as only two of the seven available water vapor channels are independent (Löhnert et al., 2009). Adding the information from 9 elevation scans performed at low angles to the standard zenith observations for the opaque oxygen complex allows to obtain temperature profiles with a higher spatial resolution and an accuracy of less than 1 K below around 1.5 km (Crewell and Löhnert, 2007).

Discharge measurements for the six subcatchments (Ach-OBN, Ach-OBH, Am-OAG, Am-PEI, Am-WM, Rt-RST) evaluated in this study are obtained from the online archive of the Bavarian Environmental Agency (LfU, https://www.gkd.bayern.de).

2.3 The WRF-Hydro modeling system

The Weather Research and Forecast modeling system (WRF, Skamarock and Klemp, 2008) is a common, community developed tool for the simulation of local area to global tropospheric dynamics and their interaction with the land surface. Applications range from short-term regional forecast to long-term continental climate studies with spatial resolutions of a few tens of meters with large-eddy simulations to several kilometers. WRF-Hydro (Gochis et al., 2016) augments WRF with respect to lateral hydrological processes at and below the land surface. It adds a surface storage layer where infiltration excess water is stored and subsequently routed according to the topographic gradient once the retention depth becomes exceeded. This is different to WRF where the infiltration excess depicts a sink term. Thus, in WRF-Hydro the surface water can infiltrate gradually, potentially leading to increased soil moisture. Gradient based routing can also be activated for saturated soil layers, and in case of oversaturation the water will exfiltrate to the surface where it enters the surface storage body and routing process. WRF-Hydro is connected with the planetary boundary layer in the same way as WRF, the lateral water transport at and below the surface is the crucial difference. Further hydrological processes that are implemented in WRF-Hydro without feedback to the

atmosphere are baseflow generation and channel routing. The model has two operation modes, stand alone, driven by gridded (pesudo-) observations (one-way coupled) or fully coupled with the dynamic atmopheric model WRF (fully-coupled).

The one-way coupled WRF-Hydro system (i.e., separate computations for atmosphere and hydrology without upward feedback) had been successfully applied for short-term forecasting (Yucel et al., 2015) and long-term hindcasting (Li et al., 2017) and was furthermore selected as the core component of the United States National Water Model (NWM, e.g., Cohen et al., 2018, http://water.noaa.gov/about/nwm). Fully, two-way coupled applications found reasonable performance for monthly scale discharge simulations for the Sissili (Arnault et al., 2016b) and the Tono (Naabil et al., 2017) basins in West Africa. Reasonable results were also achieved on a daily base for the Tana river in Kenya (Kerandi et al., 2017). In an ensemble study with the fully coupled WRF-Hydro model, encompassing six catchments in southern Germany, Rummler et al. (2018) found that simulated and observed flow exceeding percentiles on an hourly basis were in good agreement for a three months summer period in 2005. Comparison studies with respect to WRF showed slightly improved precipitation skills with WRF-Hydro for the Crati region in southern Italy (Senatore et al., 2015) and for Israel and the eastern Mediterranean (Givati et al., 2016).

WRF-Hydro provides good capability for studying the coupled land–atmospheric boundary system from catchment to continental scale regions. Although many of the recent studies focus on classic precipitation and discharge simulation performance, the ability of the fully-coupled model system to improve physical realism for water and energy budgets across compartments becomes increasingly important and is therefore of central interest in this study.

2.4 Model setup and calibration

2.4.1 Modeling chain

The study analyzes the impact of coupling hydrological processes to the regional atmospheric modeling system WRF with respect to water and energy exchange at the land surface–atmospheric boundary layer interface. Lateral flow of infiltration excess, as well as river inflow and routing are addressed by the WRF-Hydro extension. Several parameters of WRF-Hydro influence the land surface water redistribution and thus the hydrographs and require therefore thorough calibration.

To finally come up with a fully coupled WRF-Hydro setup several intermediate steps are required that involve different components of the modeling system. As outlined in Fig. 2 we build a modeling chain with the items WRF-Hydro Standalone (WRF-H_SA), WRF Standalone (WRF_SA), and WRF-Hydro Fully-Coupled (WRF-H_FC). WRF-H_SA refers to the hydrologically extended land-surface model that is not coupled to an atmospheric model and gets its driving data from gridded (pseudo-) observations. WRF Standalone (WRF_SA) is the classic version of WRF that has no hydrological extension and that is driven by data from a global circulation model. WRF-H_FC extends WRF_SA with the hydrological implementations of WRF-H_SA.

Of the 8 driving variables required by WRF-H_SA, only interpolated precipitation is available from observations with adequate coverage. Thus, the remaining input variables (temperature, humidity, wind, radiation) are taken from a preceding standalone WRF (WRF-ARW 3.7) simulation. The modeling chain encompasses the following four steps: 1) a classic standalone WRF run (2015-04-01–2016-10-31) with standard LSM parameters to derive the driving variables required by 2) the stan-

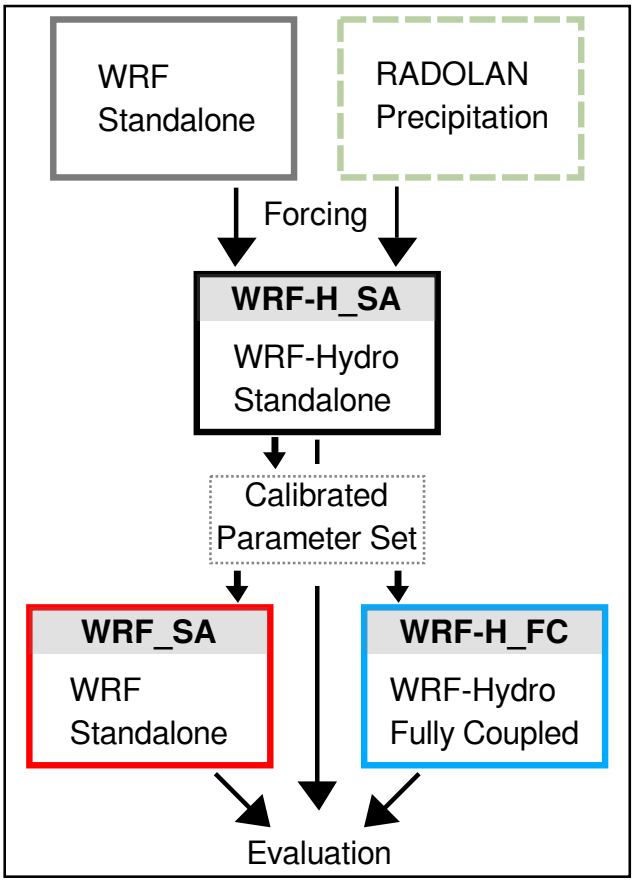

Figure 2. Modeling components and workflow. RADOLAN-RW is an hourly precipitation observation product of the German Weather Service with 1 km^2 grid resolution.

dalone WRF-Hydro simulations (WRF-H_SA, 2015-04-01–2015-07-31 and 2016-04-15–2016-10-31) that ingest also gridded observed precipitation (RADOLAN, Bartels et al., 2004; Winterrath et al., 2012), 3) a fully coupled WRF-Hydro simulation (WRF-H_FC 2016-04-15–2016-10-31) using calibrated parameters from WRF-H_SA, and 4) a rerun of the classic standalone WRF (WRF_SA, 2016-04-15–2016-10-31) with the same parameter set obtained from WRF-H_SA. Finally, this leads to a commensurable set of simulations, coupled versus uncoupled. Consequently, WRF_SA does not represent an optimized setup of a classic WRF stand alone model. Therefore, it is possible that with other parameter combinations for WRF_SA even better performance could be achieved. However, tuning WRF_SA is not easily possible because it does not feature the simulation of discharge and other point observations are sparse and represent different scales and are thus only suitable for the evaluation of simulation results.

The computational demand was about 0.021 million core hrs for the WRF_SA simulations, 0.32 million core hrs for the WRF-H_SA calibration runs and 0.042 million core hrs for WRF-H_FC on a 2.3 GHz Intel Haswell system.

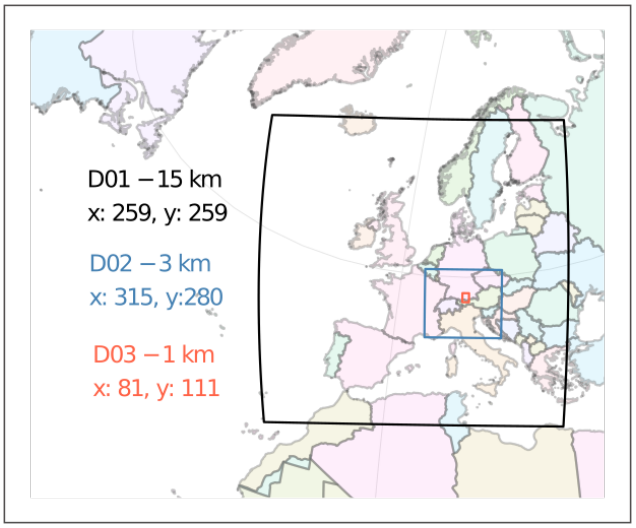

Figure 3. Domain nesting configuration for WRF_SA and WRF-H_FC simulations.

2.4.2 Space and time

Figure 3 visualizes the domain and nesting configuration for the WRF_SA and WRF-H_FC simulations. A telescoping configuration with 3 nests is employed. The horizontal resolutions are 15 by 15, 3 by 3, and 1 by 1 km for domain 1, 2, and 3, respectively. The finest domain extends from the city of Munich in the northeast to the mountain valleys of Inn and Lech in the southwest. For all domains, the number of vertical levels is 51. Model top is defined at 10 hPa. The WRF-H_SA and WRF-H_FC simulations cover the period 2016-04-15 to 2016-10-31 including a half month for model spin-up. The starting date corresponds to snow free conditions for most of domain 3. The spin-up strategy avoids the uncertain simulation of the snow storage dynamics for the winter season. For the surface variables in WRF-Hydro we consider a 15 day spin-up to be sufficient. The initial (2016-04-15) soil moisture fields for both WRF_SA and WRF-H_FC are taken from the last WRF-H_SA simulation timestep (2016-10-31). We assume that this 6 month spin-up period is sufficient to come up with reasonable starting conditions for a commensurable set of simulations. The model runs are performed continuously with none of the variables being reinitialized in between. Lateral surface water flow processes (i.e., overland and channel routing) in WRF-H_SA and WRF-H_FC are computed on a 100 by 100 m grid with the extent being identical to that of domain 3. The integration timesteps for the atmospheric part (WRF including Noah-MP LSM) are 60, 12, and 4 s for domains 1, 2, and 3, respectively. The hydrological routines are called at hourly intervals.

2.4.3 Model physics

The WRF physics parameterization for the selected domains is listed in Tab. 1. Cumulus parameterization is only activated for the outermost domain (1), while explicit convection is chosen for the finer grids (domain 2 and 3), according to Skamarock et al. (2008).

Table 1. Selected WRF_SA and WRF-H_FC model physical schemes.

Physics categories	Selected scheme	Reference
Microphysics	Thompson	Thompson et al. (2008)
Cumulus parameterization	Kain-Fritsch*	Kain (2004)
Planetary boundary layer	QNSE	Sukoriansky et al. (2005)
Land surface model	Noah-MP	Niu et al. (2011)
Longwave radiation	RRTMG	Iacono et al. (2008)
Shortwave radiation	RRTMG	Iacono et al. (2008)

* Only for the outermost domain

For uncoupled and coupled simulations, Noah-MP (Niu et al., 2011) is used as the land surface model. For Noah-MP the selected configuration deviates from the default setup as follows: the Community Land Model (CLM) method is selected for stomatal resistance computation, the Schaake et al. (1996) method is used to determine infiltration and drainage (similar to classic Noah-LSM), the two stream radiation transfer applied to vegetated fraction (option 3) is chosen, and the dynamic vegetation option (Dickinson et al., 1998) is enabled.

The static data is adopted from the standard WRF geographic dataset. The landcover for domain one and two is based on USGS classification and includes lakes. For the innermost domain the landcover information is based on the CORINE (Büttner, 2014) dataset of the European Union, reclassified according to the USGS classes. The elevation data for the 100 m grid is derived from the ASTER global digital elevation model (ASTER GDEM, version 2).

Noah-MP provides different options for the computation of groundwater discharge and surface runoff (infiltration excess) but only WRF-Hydro enables the simulation of lateral hydrological processes such as the overland routing of surface runoff and channel (discharge) routing. A short description of the most relevant model details for this study is provided below. Further information about technical features and standard model physics options are given in Gochis et al. (2016), while Section 2.4.4 describes the specific improvements made to the original model in order to fit with the specific features of the complex topography of this study area.

In the WRF-Hydro modeling system only subsurface and surface overland flow routing are allowed to directly affect atmosphere dynamics (i.e., only these processes are fully coupled). After every LSM loop, a sub-grid disaggregation loop (Gochis and Chen, 2003) is run prior to the routing of saturated subsurface and surface water, in order to achieve the desired spatial refinement (from 1 km to 100 m) for the two state variables infiltration excess, and soil moisture content. At this stage, linear sub-grid weighting factors are assigned for preserving the sub-grid soil moisture and infiltration excess spatial variability structures from one model timestep to the next. Then, subsurface lateral flow is calculated, using the method suggested by Wigmosta et al. (1994) and Wigmosta and Lettenmaier (1999) within the Distributed Hydrology Soil Vegetation Model (DHSVM). The water table depth is calculated according to the depth of the top of the highest (i.e., nearest to the surface) saturated layer. Finally, overland flow routing also accounts for possible exfiltration from fully saturated grid cells and is achieved through a

fully-unsteady, explicit, finite difference, diffusive wave approach similar to that of Julien et al. (1995) and Ogden (1997). In this study the steepest descent method is used with a timestep of 6 s. After the execution of the routing schemes, the fine grid values are aggregated back to the native land surface model grid.

Concerning the one-way coupled processes modeled in the WRF-Hydro system (i.e., no feedback with the atmosphere), in this study, the channel flow and baseflow modules are used. Specifically, channel flow routing is performed through an explicit, one-dimensional, variable timestepping diffusive wave formulation. Overland flow discharging into the stream channel occurs when the ponded water depth of specific grid cells, assigned to a predefined stream channel network, exceeds a fixed retention depth. The channel network has a trapezoidal geometry, depending on the Strahler stream order functions. Currently no overbank flow is simulated. Baseflow to the stream network is represented through a simple bucket model which uses an exponential equation to achieve the bucket discharge as a function of a conceptual water depth in the bucket. Several baseflow sub-basins (i.e., several conceptual buckets) can be specified within a watershed, but since an empirical equation is used, its parameters need to be estimated for each of the sub-basins. The baseflow model is linked to WRF-Hydro through the deep drainage discharge from the land surface soil column. Estimated baseflow discharged from the bucket model is then combined with lateral inflow from overland flow and is input directly into the stream network as a part of the stream inflow. Total sub-basin baseflow flux to the stream network is equally distributed among all channel pixels within the sub-basin.

The reservoir module (retention of channel flow in lakes and reservoirs) is disabled for the simulations of this study because only one gauge (Ach-OBH) would be slightly affected by a 7.66 km^2 lake (Staffelsee) but the number of calibration parameters and thus calibration runs would considerably increase. Since the lake evaporation is explicitly considered by the Noah-MP land surface model we do not assume any impact on the land surface–boundary layer exchange by this simplification.

2.4.4 Changes with respect to the original WRF-Hydro model

The model, as applied in this study, differs from the version 3 of WRF-Hydro with respect to soil layer representation and model timesteps. Large parts of the modeling domain and the considered river catchments exhibit mountainous terrain with steep slopes covered by shallow soil layers. Here, the model's general assumption of two meter soil thickness (or depth to bedrock) does not hold true as it may lead to overestimated retention of infiltrating water. Therefore, in this study, the soil layer definition is changed from a domain uniform to a grid point based representation and soil layer thicknesses are set to the Noah-MP standard (0.1, 0.3, 0.6, 1 m) distribution for hillslopes below 50 % and to more shallow values (0.05, 0.05, 0.1, 0.1 m) for all slopes >50 %. The 50 % threshold leads to a realistic discriminability of valley bottoms and hillslopes. In addition, the infiltration (REFKDT) and percolation (SLOPE) parameter implementation is changed from domain-wide uniform values to subcatchment-wise distributed (lumped). Another important change is made with respect to the timestep configuration of WRF-H_SA and WRF-H_FC simulations. As pointed out in Senatore et al. (2015), differing intervals used for temporal integration of the Noah-LSM lead to inconsistent amounts for the soil water fluxes. The issue is caused by numeric effects, when timesteps become very small (1 km WRF requires about 4 s) in fully coupled WRF-Hydro simulations with high resolutions. To eliminate the problem and to make all simulations comparable, the hydrology part (subgrid) in WRF-H_FC is only called at an hourly timestep, similar to that of WRF-H_SA. In WRF-H_FC the flux variables of domain 3 are therefore cumulated in between the

(hourly) calls of the hydrological routines and on the other hand, overland routing output (surface head) is returned to domain 3, equally distributed over the LSM timesteps (4 s). The subsurface routing option for saturated soil layers is deactivated in this study as full saturation takes rarely place in the region and because the module would require to extend the calibration by two more parameters.

2.4.5 Driving data

Atmospheric boundary conditions for the outer domain of the WRF_SA and the WRF-H_FC simulations are derived from the ERA Interim reanalysis (Dee et al., 2011) with 0.75 ° horizontal grid spacing, 37 pressure levels from 1000 to 1 hPa, and 6 hours temporal resolution. Forcing data for WRF-H_SA are taken from the standard WRF simulation output for domain 3. The variables comprise near surface air temperature, humidity, wind, surface pressure, short- and longwave downward radiation. Since precipitation from WRF simulation is typically biased and dislocated, an observational product of the German weather service (RADOLAN, Bartels et al., 2004; Winterrath et al., 2012) is used for substitution. It combines gauge and rain radar information and is available with an hourly timestep and 1 km^2 resolution.

2.4.6 Calibration

Different approaches for the calibration of the WRF_SA model have been followed in previous works, all based on the comparison of the observed hydrographs. Yucel et al. (2015) adopted a stepwise approach, where the parameters controlling the total water volume were first calibrated (namely, the infiltration factor, REFKDT, and the surface retention depth, RETDEPRT), and then the parameters controlling the hydrograph shape (namely, the surface roughness, OVROUGHRT, and the channel Manning roughness, MANN). Li et al. (2017), Naabil et al. (2017), Kerandi et al. (2017) and Senatore et al. (2015) followed a similar approach. Specifically, the latter added in the first calibration step the parameter that governs deep drainage (SLOPE) and in the second step the saturated soil lateral conductivity and the bucket outflow exponent (EXPON). Furthermore, to refine calibration they introduced an automated procedure based on the *Parameter Estimation and Uncertainty Analysis software* (PEST, Doherty, 1994). Arnault et al. (2016b) mainly focused on the REFKDT and, secondarily, on the MANN parameters. Finally, Silver et al. (2017) proposed a satellite-based approach for arid (bare soil) regions aimed at calibrating topographic slope, saturated hydraulic conductivity and infiltration parameters, based on physical soil properties and not depending on observed runoff.

Calibrating a complex hydrological model with a large number of parameters by means of only river discharge can be very problematic, particularly because of the known problem of equifinality (Hornberger and Spear, 1981; Beven, 2006; Beven and Binley, 2014). Several approaches are adopted to reduce or control this problem, particularly challenging for the emerging fully distributed paradigm in hydrology (Beven and Binley, 1992; Beven, 2001; Kelleher et al., 2017), either constraining the parameter set by means of various strategies (e.g., Cervarolo et al., 2010) and/or incorporating different observations than discharge in the calibration process (e.g., Thyer et al., 2004; Stisen et al., 2011; Graeff et al., 2012; Corbari and Mancini, 2014; Larsen et al., 2016b; Soltani et al., 2019). A fully-coupled atmospheric–hydrological approach further increases the degrees of freedom of the model, making the issue even more complex. In this study, while the calibration of the hydrological model is

Table 2. List of parameters used for WRF-H_SA calibration.

Parameter	Abbreviation	Sensitivity	Compartment
Surface roughness scaling factor	ovrgh	yes	overland routing
Retention depth scaling factor	retdp	yes	overland routing
Infiltration coefficient	refkdt	yes	LSM
Free drainage coefficient	slope	yes	LSM
Bucket storage height	zmax	yes	baseflow model
Bucket storage initial water	zinit	no	baseflow model
Bucket outflow coefficient	coeff	yes	baseflow model
Bucket outflow exponent	expon	yes	baseflow model

performed offline, accounting only for discharges from several cross river sections, the effect of the resulting parameter set is evaluated considering soil, surface (both in terms of vegetation and hydrology) and atmosphere compartments all together with their reciprocal interactions. Further research will focus on more thorough analysis of equifinality issues in two-way coupled hydrometeorological models.

After several preliminary runs, where the model sensitivity to all the parameters involved in literature calibration procedures is tested, the WRF-H_SA model calibration follows also a two step approach, but in a different sense with respect to Yucel et al. (2015). First, the *Latin-Hypercube One-factor-At-a-Time* (LH-OAT, Van Griensven et al., 2006) method is used to determine the sensitivity of a set of 8 selected parameters on sub-basin river discharge, but also to obtain a starting configuration for the automated parameter optimization. The first step includes some iterations to find an optimal threshold value for the delineation of shallow soils on the steep mountain slopes and deeper soils in the plains (see also 2.4.4). In the next step, 7 sensitive parameters are optimized for the 6 different sub-basin outlets (Fig. 1b) using PEST (Doherty, 1994). Table 2 gives an overview of the parameters and their relevance.

The calibration procedure is adopted for the different subcatchments in cascade, starting from upstream (i.e., parameters are first calibrated for Am-OAG, Ach-OBN and Rt-RST, then for Am-PEI and Ach-OBH and, finally, for Am-WM, see Fig. 1b). For LH_OAT the goodness of fit is determined using the volumetric efficiency (VE, Criss and Winston, 2008):

$$VE = 1 - \frac{\sum |Q_{obs} - Q_{sim}|}{\sum Q_{obs}}$$

with Q_{Obs} and Q_{sim} denoting the observed and simulated discharge in $m^3 s^{-1}$, and the Nash-Sutcliffe Efficiency (NSE). PEST optimization relies on an objective function given by the sum of squared deviations between model-generated streamflow and observations. Table 3 lists the subcatchment-wise calibrated parameters. The optimum slope gradient for the delineation of shallow and deep soil regions is found to be 50% (22.5°). Accounting for the shallow mountain slopes considerably improves the hydrographs for Am-OAG and Am-PEI, as it increases underestimated peaks and decreases overestimated retention. A

Table 3. Calibration parameters for the six subcatchments of Ammer (Am), Ach, and Rott (Rt), at the gauges Oberammergau (OAG), Peißenberg (PEI), Weilheim (WM), Obernach (OBN), Oberhausen (OBH), and Raisting (RST)

Basin	Am-OAG	Am-PEI	Am-WM	Ach-OBN	Ach-OBH	Rt-RST
coeff	5.00	11.00	2.41	3.10	4.27	0.010
expon	3.00	11.00	0.01	1.50	1.58	0.631
zmax	1.00	1.41	58	1.00	29	1.0
refkdt	0.74	0.028	0.196	0.061	3.73	0.106
slope	0.25	0.20	0.14	0.05	0.06	0.025
ovrgh	65.2	15.60	37.6	40.9	5.88	26.30
retdp	0.73	5.00	0.64	3.16	3.82	0.100
VE	0.54	0.56	0.50	0.54	0.78	0.46
NSE	0.36	0.21	0.68	0.56	0.60	0.56
Shift	2.00	4.71	5.13	-	-	-
VE_shifted	0.81	0.72	0.79	-	-	-
NSE_shifted	0.64	0.64	0.62	-	-	-

comparison of the respective hydrographs and performance measures for calibration and validation period is shown in figures S1 and S2.

Since the study focuses on land atmosphere exchange, and river routing has no feedback to the LSM, the channel parameters (geometry, Manning roughness coefficient) are not further optimized with respect to peak timing.

The calibration period length of 3.5 months (2015-04-15–2015-07-31, including 14 days of spin-up) is selected as a compromise between the number of model runs (about 2000, which relates to about 0.32 MIO core hours for WRF-H_SA), required during hypercube sampling and PEST optimization, and the available computational resources.

The hydrographs for the calibration and validation periods of the WRF-H_SA runs are presented in Figures 4 and 5. The final parameter sets and goodness of fit measures are listed in Table 3. For all subcatchments, reasonable configurations could be determined. However, for the three Ammer subcatchments (OAG, PEI, WM) it was required to add constant baseflow rates of 2, 4.71, and 5.13 $m^3\ s^{-1}$ to the simulated hydrographs, respectively. Adding constant baseflow is justified by the fact that due to the glacial processes of the last ice-age, large storage bodies, that dip reversely with the surface elevation, were formed in the mountain valleys by overdeepening (Seiler, 1979; Frank, 1979). Further downstream, towards the opening of the valley, where the aquicludes reach towards the surface, springs are abundant. Channel inflow from such long-term storage cannot be realized by WRF-Hydro's conceptual bucket scheme. It would require the implementation of a more sophisticated groundwater model. The underestimation of long-term baseflow by the model has no influence on the land surface–atmosphere exchange as there is no interaction of the channel routing with the LSM. The amounts for constant baseflow are derived manually after the PEST parameter optimization, so that the recession curves of the simulations agree well with the observations. For the Ach and Rott

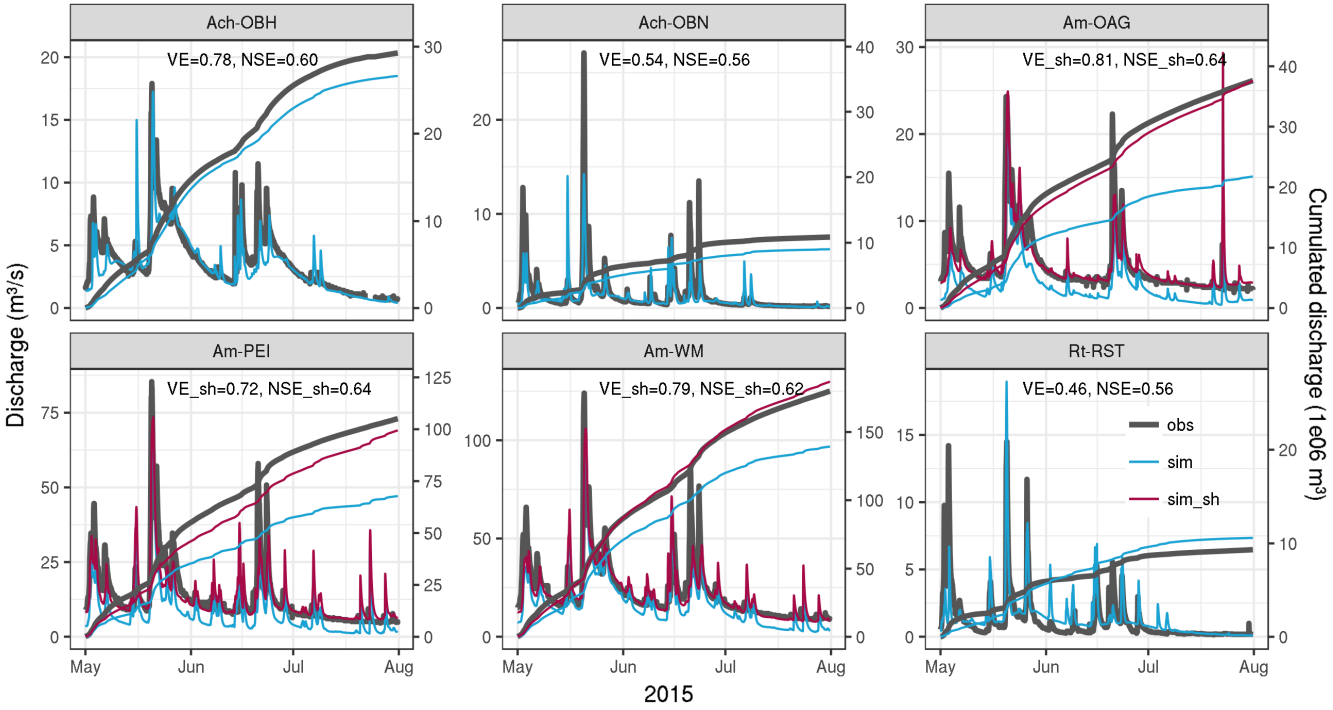

Figure 4. Subcatchment hydrographs for calibration period (2015-05-01–2015-07-31). Standard WRF-H_SA model output is printed in blue. Shifted (sh) hydrographs are shown in red. Shift amounts are listed in Tab. 3.

subcatchments channel baseflow is related with shallow aquifers with shorter residence times which should be solely captured by the model.

Focusing on the values of the calibrated parameters, it can be observed that the surface infiltration parameter REFKDT, as compared to the standard settings in WRF and Noah-MP (e.g., nominal range of 0.5-5.0, according to Niu (2011); 0.1-0.4, according to Lahmers et al. (2019)), is rather low (therefore allowing lower infiltration) for all subcatchments except Ach-OBH. The associated LSM surface runoff scaling parameter REFDK is globally set to $2e^{-06}$ as smaller values would have decreased infiltration to even smaller amounts. Also, the percolation parameter SLOPE was mainly reduced as compared to the standard values (0.1-1.0, according to Niu et al., 2011b), meaning that a relatively limited portion of former infiltration excess water needed to be transferred to the bucket-storage to assure good performance for the simulated baseflow. As for the surface overland roughness scaling factors, they generally increase remarkably the standard value of 1.0, contributing to the increase of the hydrograph's lag time and the relative reduction of the peak discharge. The retention depth scaling factors, instead, are much closer to the standard value of 1, varying slowly both the total volumes of the hydrographs and the lag time of the initial response of the catchment to rainfall. The bucket scheme parameters should be evaluated considering their mutual influence on the model exponential equation. In general, the higher the ZMAX value, the slower the response time of the bucket, and the higher the COEFF value, the higher the potential contribution of the bucket model to the total runoff. From

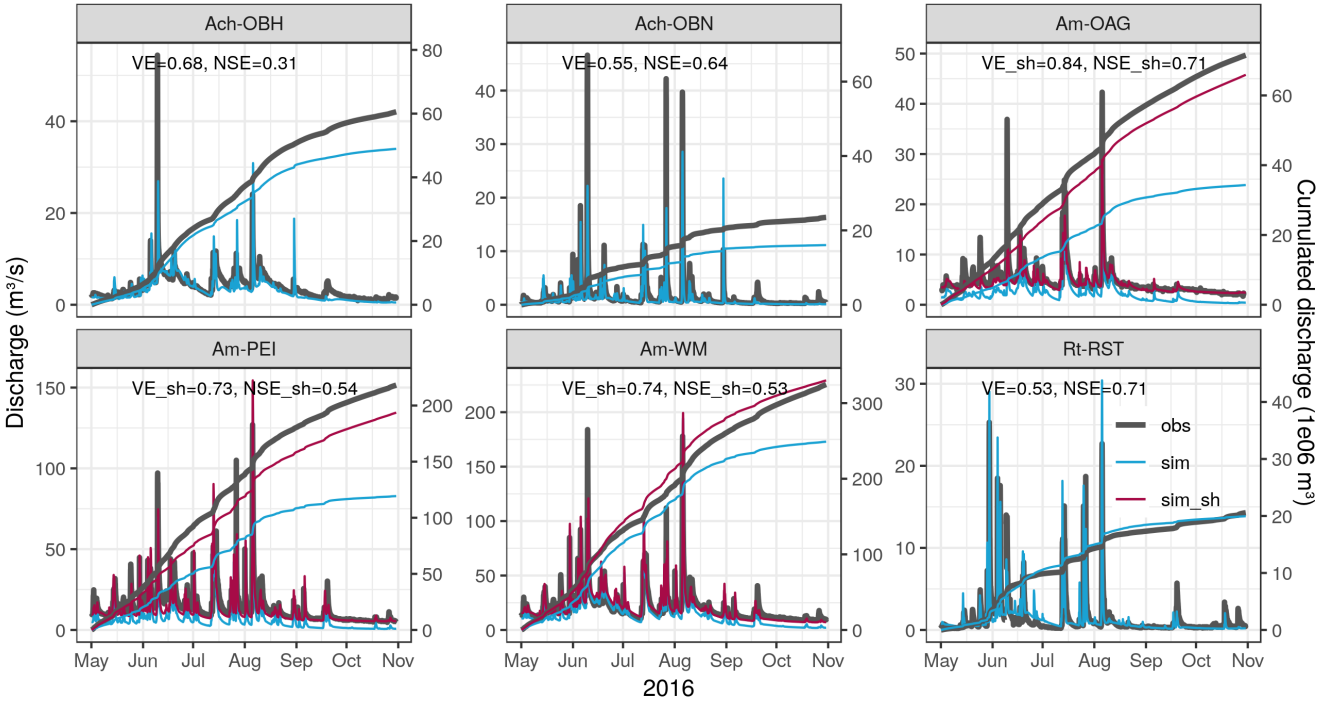

Figure 5. Subcatchment hydrographs for validation period (2016-05-01–2016-10-31). Standard WRF-H_SA model output is printed in blue. Shifted (sh) hydrographs are shown in red. Shift amounts are listed in Tab. 3.

this point of view, the most reactive subcatchment is Am-PEI. In this subcatchment, REFKDT is rather small and therefore the contribution by the bucket needs to be quicker also because of the rather large subcatchment site. However, this bucket's behavior cannot be appreciated by looking at the hydrographs in Fig. 4, both because in the simulations the values of the bucket storage height are never close to ZMAX, and because the resulting discharge in the graph also depends on the contribution of the upstream catchment Am-OAG. During the PEST automated calibration, it occurred that some parameters hit a boundary limit of the calibration range, previously set starting from the results obtained with the LH-OAT method. E.g., the optimized ZMAX values hit the lower boundary for the three upstream catchments. In such cases the constraints were relaxed, allowing the parameters to exceed the previous limits, but if negligible or even null improvements were found, we preferred to come back to the previous borders. Finally, it is noteworthy to highlight that, since the study focuses on land-atmosphere exchange, and river routing has no feedback to the LSM, the channel parameters (geometry, roughness coefficient) are not further optimized concerning peak timing.

The calibration performed in the spring/summer of 2015 is validated over the period 2016-05-01–2016-10-31 (Fig. 5). The performance statistics for the validation period are comparable to those of the calibration period and in the cases of Ach-OBN, Am-OAG, and Rt-RST even improved. For Ach-OBH, as expected due to the disabling of the reservoir option, the buffering effect of the lake cannot be reproduced by the model, thus leading to an overestimation of most of the peak values. However,

with respect to the overall discharge in the Ammer catchment, these cut-off peaks are rather small. The performance for Am-WM is lower for both 2015 and 2016, as it aggregates the mismatches of all the upstream subcatchments.

The results of the calibration are in line with other hydrological modeling studies for the Ammer catchment. Ludwig and Mauser (2000) implemented TOPMODEL (Beven et al., 1984) into a SVAT model framework and yielded a NSE of 0.92 for one year simulation on a daily basis for the gauge Fischen (nearby Am-WM). Marx (2007) achieved NSE performances of 0.2 (Am-OAG), 0.42 (Am-PEI), 0.75 (Am-WM), 0.68 (Ach-OBN), and 0.18 (Ach-OBH), using WaSiM-ETH (Schulla and Jasper, 2007) for the year 2001. Rummler et al. (2018) obtained a NSE of 0.91 with the WRF-Hydro standalone model for Am-PEI, for a 3 months simulation of a major flood event in 2005.

The commonly favored lumped calibration of WRF-Hydro seems to be rather limited, concerning the transferability of parameter sets among subcatchments and with respect to the numerical efficiency for automated calibration. Especially for complex terrain, e.g., as presented by this study, the distribution of discharge gauges does not agree with landscape units. Therefore, the lumped parameter sets have to union quite diverse subcatchment conditions which may lead to unrealistic spatial representations of the physical properties they represent. Thus, for further studies, it is recommended to find parameter sets that are bound to landscape characteristic, such as relief, landcover type and soil features (e.g., Hundecha and Bárdossy, 2004; Samaniego et al., 2010; Rakovec et al., 2016; Silver et al., 2017), which also contribute to reduce the equifinality problem (Kelleher et al., 2017).

3 Results and discussion

The following section evaluates and discusses the simulations of the standalone WRF (WRF_SA) and the fully coupled WRF-Hydro (WRF-H_FC). In the first part, based on the TERENO Pre-Alpine stations, the energy fluxes at the land–atmosphere boundary are analyzed, in particular radiation, heat fluxes, and near surface air temperature, evapotranspiration and soil moisture. The second part compares modeled and observed atmospheric boundary layer profiles for the DE-Fen site. The third part deals with the subcatchment aggregated water budgets and looks at the differences in the temporal evolution of simulated soil moisture patterns.

3.1 Model evaluation for TERENO Pre-Alpine stations

3.1.1 Radiation

The evaluation of WRF_SA and WRF-H_FC simulations with observations from TERENO Pre-Alpine focuses on the radiation input, its partitioning into water and energy fluxes at the land surface and on the near surface atmospheric and subsurface states. Fig. 6 shows the mean diurnal cycles of simulated and observed downward surface shortwave radiation for three TERENO Pre-Alpine sites for the period June to October 2016. For all locations, the simulations overestimate radiation from sunrise to sunset with similar magnitude. While correlations for the hourly values are high (r^2: DE-Fen 0.76, DE-RbW 0.73, DE-Gwg 0.66), the mean errors (ME W/m^2: DE-Fen 61, DE-RbW 57, DE-Gwg 78) reveal considerable bias. Also the mean absolute errors

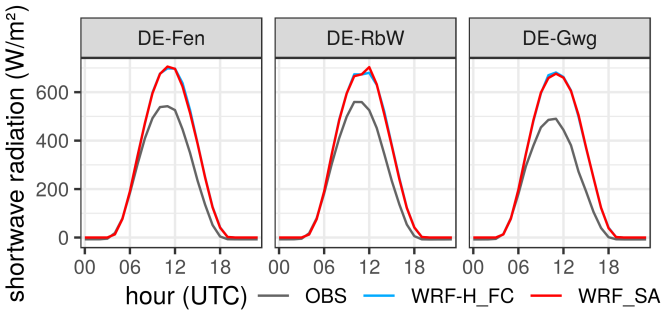

Figure 6. Mean diurnal cycles of simulated and observed downward surface shortwave radiation for June to October 2016.

show substantial scatter (MAE W/m^2: DE-Fen 89, DE-RbW 91, DE-Gwg 107). The overestimation of summer shortwave radiation for Central Europe with WRF has also been documented by other studies and is usually related to underestimated cloud cover, especially in the mid troposphere where convection is active (García-Díez et al., 2015; Katragkou et al., 2015). The increased bias for DE-Gwg could be related to local shading due to topography in this narrow Alpine valley and because of higher convective activity in this mountain region. The comparison of WRF_SA and WRF-H_FC simulation does not yield considerable differences.

The results for downward longwave radiation are given in Fig. 7. The negative biases for the different locations (ME W/m^2:

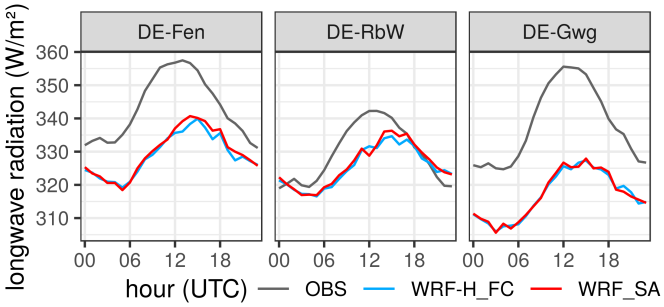

Figure 7. Mean diurnal cycles of simulated and observed downward longwave radiation for June to October 2016.

DE-Fen -14, DE-RbW -4, DE-Gwg -21) correspond with the above suggested cloud cover underestimation. With -1 to -6 %, the relative deviations are rather small. Again, the differences between standalone and coupled model are nominal. Similar overestimation is obtained for total absorbed shortwave radiation (Tab. S1). Thus, for the TERENO Pre-Alpine locations it can be stated that shortwave radiation input to the land surface is overestimated with 33 to 56 % by both standalone and coupled model.

3.1.2 Heat fluxes and evapotranspiration

The diurnal cycles of latent and sensible heat fluxes are presented in Figs. 8 and 9 for three different grassland sites. Converted evapotranspiration rates are plotted alongside in Fig. 8. The analysis is split by month from June to October 2016, to provide insight about the temporal variations. The observations comprise data from flux-towers and lysimeters. Their spread can be seen as a measure of uncertainty. The measurements for the flux-tower at DE-RbW are missing from 25th September to end of October. The coupled WRF-H_FC simulation exhibits increased latent and decreased sensible heat fluxes for the DE-Fen

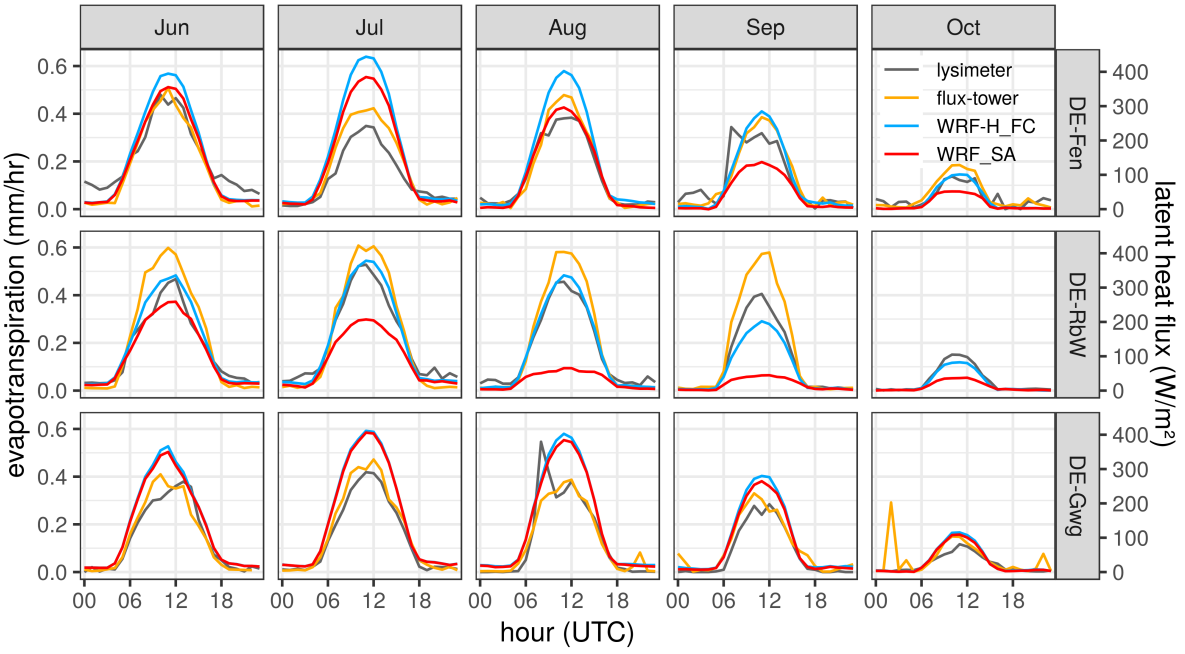

Figure 8. Mean diurnal cycles of simulated and observed latent heat flux / evapotranspiration for the months June to October 2016 at different TERENO Pre-Alpine sites.

and DE-RbW sites, whereas WRF_SA and WRF-H_FC are very similar for the Alpine valley location (DE-Gwg). In the case of sensible heat flux, the hydrologically enhanced model (WRF-H_FC) outperforms or is equal to the standalone simulation (WRF_SA). For latent heat / evapotranspiration the mean diurnal fluxes are overestimated for DE-Fen for June to August by either both models or the coupled run. A constant positive bias is also found for DE-Gwg (except for Oct.). Tabs. 4 and 6 list the performance measures for latent and sensible heat versus flux-tower observations for the period June to October 2016, Tab. 5 provides the measures for evapotranspiration versus lysimeter data. For latent heat and likewise evaporation, correlation improves considerably for DE-Fen and DE-RbW with the coupled model but ME and MAE deteriorate for DE-Fen and DE-Gwg. Overall improvement for WRF-H_FC with respect to the observations is only obtained for DE-RbW. In general, the

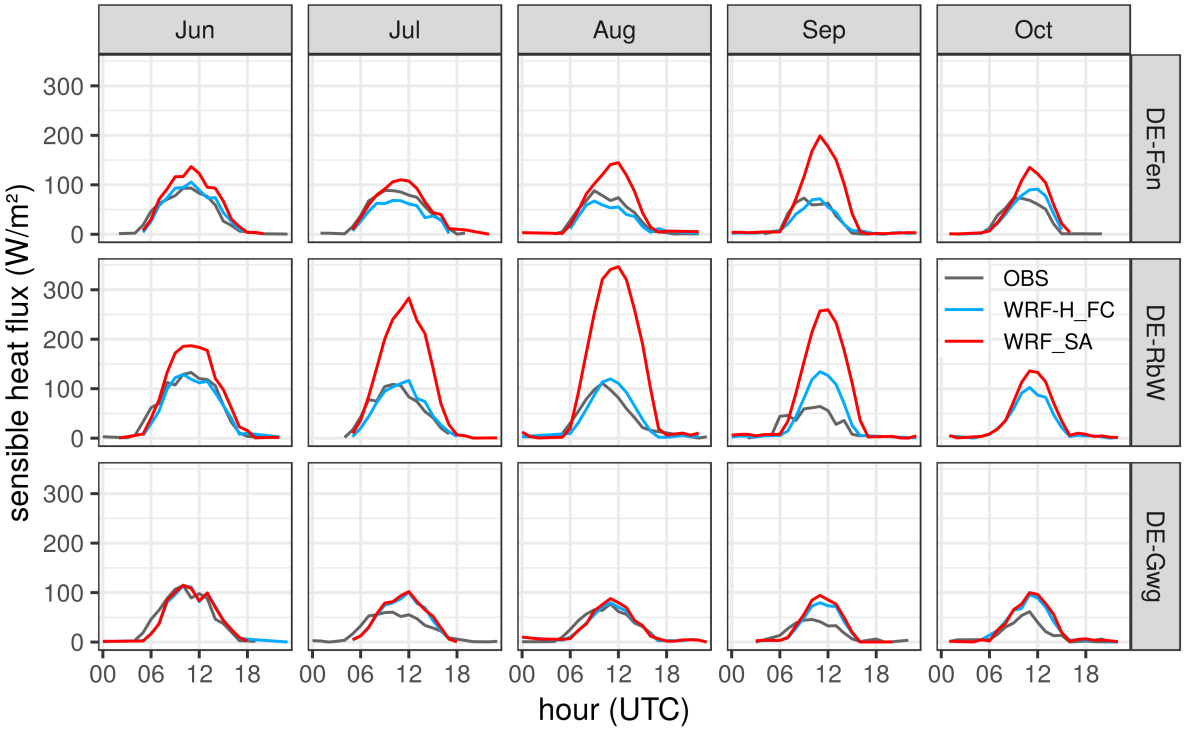

Figure 9. Mean diurnal cycles of simulated and observed sensible heat flux for the months June to October 2016 at different TERENO Pre-Alpine sites.

Table 4. Performance measures for latent heat flux (W m^{-2}), simulations vs. flux-tower observations, June to October 2016.

Station	Model	r^2	ME	MAE
DE-Fen	WRF_SA	0.59	3.96	65.24
	WRF-H_FC	0.67	38.33	67.91
DE-RbW	WRF_SA	0.44	−73.58	93.99
	WRF-H_FC	0.77	−6.59	54.19
DE-Gwg	WRF_SA	0.60	48.05	75.84
	WRF-H_FC	0.59	53.42	80.12

simulations are in better agreement with the flux-tower data than with those of the lysimeters. This could be related to the difference in spatial support of the measurements.

Table 5. Performance measures for evapotranspiration (mm hr^{-1}), simulation vs. lysimeter observation, June to October 2016.

Station	Model	r^2	ME	MAE
DE-Fen	WRF_SA	0.32	0.00	0.09
	WRF-H_FC	0.34	0.04	0.10
DE-RbW	WRF_SA	0.25	−0.07	0.11
	WRF-H_FC	0.44	0.00	0.09
DE-Gwg	WRF_SA	0.26	0.04	0.09
	WRF-H_FC	0.25	0.05	0.10

For sensible heat, the coupled run yields improved performance for DE-Fen and DE-RbW and is also in good agreement with the observations at DE-Gwg. Ground heat flux (Fig. S3) is overestimated by the models from about two hours after sunrise

Table 6. Performance measures for sensible heat flux (W m^{-2}), simulations vs. observations, June to October 2016.

Station	Model	r^2	ME	MAE
DE-Fen	WRF_SA	0.37	17.99	43.49
	WRF-H_FC	0.41	−6.46	33.67
DE-RbW	WRF_SA	0.48	46.43	64.69
	WRF-H_FC	0.59	−6.98	29.09
DE-Gwg	WRF_SA	0.23	5.18	39.85
	WRF-H_FC	0.23	2.00	38.99

until noon. From afternoon till dawn, both simulations overestimate the upward (land to atmosphere) radiative flux. WRF_SA and WRF-H_FC differ slightly for ground heat fluxes, with WRF-H_FC showing slightly increased performance with respect to the observations (Tab. S2).

3.1.3 Near surface temperature and humidity

The diurnal course of 2 m air temperature (Fig. S4) is similar for uncoupled and coupled model for June, October for all stations. The mountain peak stations (Kol, LaS) and the alpine valley station (DE-Gwg) are hardly sensitive to coupling. Prominent deviations between WRF_SA and WRF-H_FC occur for July to September at the foreland stations (DE-Fen, DE-RbW, Ber). Here, the coupled simulations between 06 and 18 UTC agree better with the observations. Nighttime values are generally overestimated by both models and coupling does not have an influence. The mean errors improve between 0.34 and

0.6 K for the foreland and between 0.11 and 0.25 K for the mountain stations whereas correlation remains identical. Slight improvement with coupling is also obtained for the MAE values (Tab. S3).

Figure 10 provides the monthly diurnal cycles for 2 m mixing ratio. The model comparison reveals higher values for the coupled model run, especially during sunshine hours (06-18 UTC). Also prominent is a peak in 2 m moisture around 1700 UTC for both models that is not as pronounced in the observations. For July to August, the coupled simulation resembles

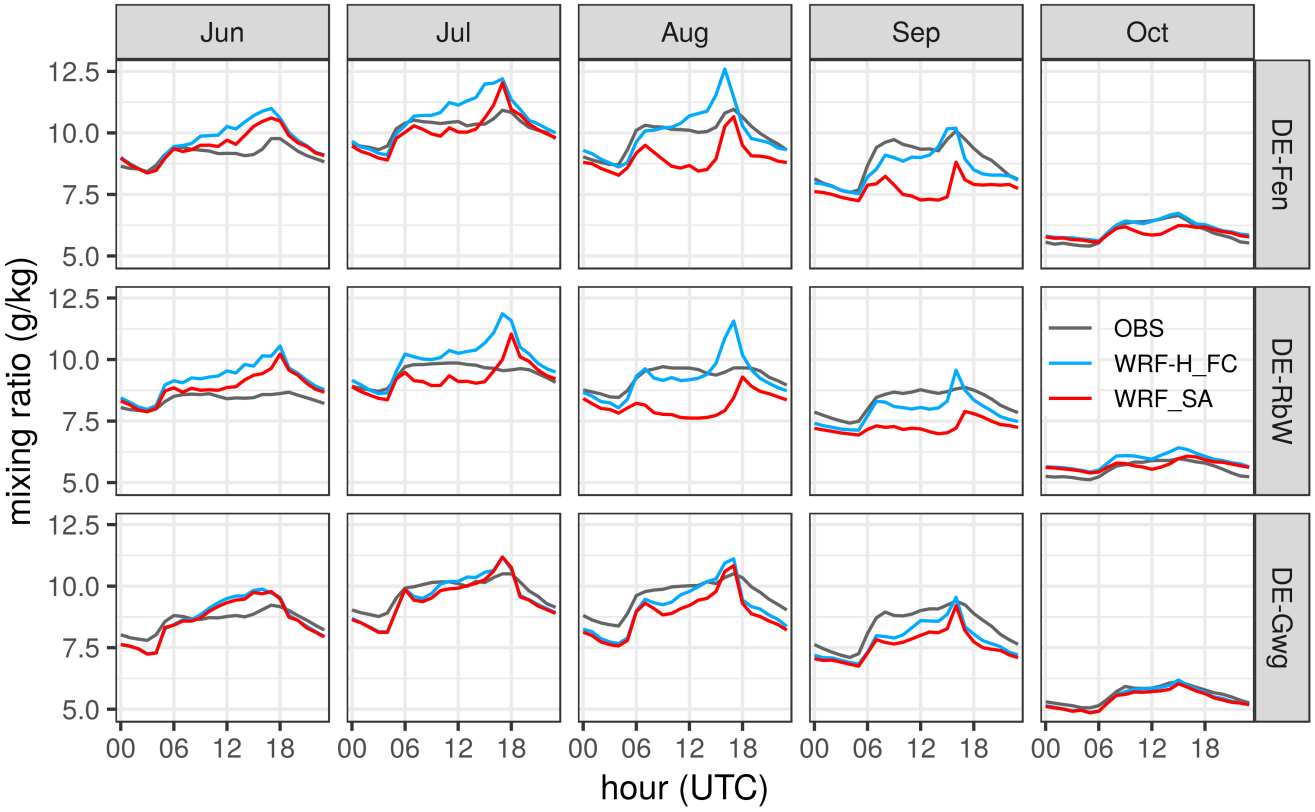

Figure 10. Mean diurnal cycles of simulated and observed 2 m mixing ratio for the months June to October 2016 at different TERENO Pre-Alpine sites.

the observations better for the morning rise in moisture concentration but towards the afternoon, the constant rise exceeds the measurements. According to the performance measures in Table 7,

the correlation increases considerably with the WRF-H_FC configuration for the DE-Fen and DE-RbW sites. Also ME and MAE are reduced. For DE-Gwg the findings are likewise, however the magnitudes are lower.

Altogether, it can be stated that the hydrologically enhanced setup (i.e., WRF-H_FC) leads to an improved representation of 2 m temperature and mixing ratio.

Table 7. Performance measures for 2 m mixing ratio (g kg^{-1}), simulations vs. observations, June to October 2016.

Station	Model	r^2	ME	MAE
DE-Fen	WRF_SA	0.57	−0.38	1.19
	WRF-H_FC	0.69	0.19	1.02
DE-RbW	WRF_SA	0.46	−0.35	1.26
	WRF-H_FC	0.62	0.25	1.08
DE-Gwg	WRF_SA	0.67	−0.39	1.03
	WRF-H_FC	0.70	−0.24	0.96

3.1.4 Soil moisture

Observed and simulated soil moisture for the DE-Fen site are presented in Figure11. The gray ribbons depict the 25-75 per-

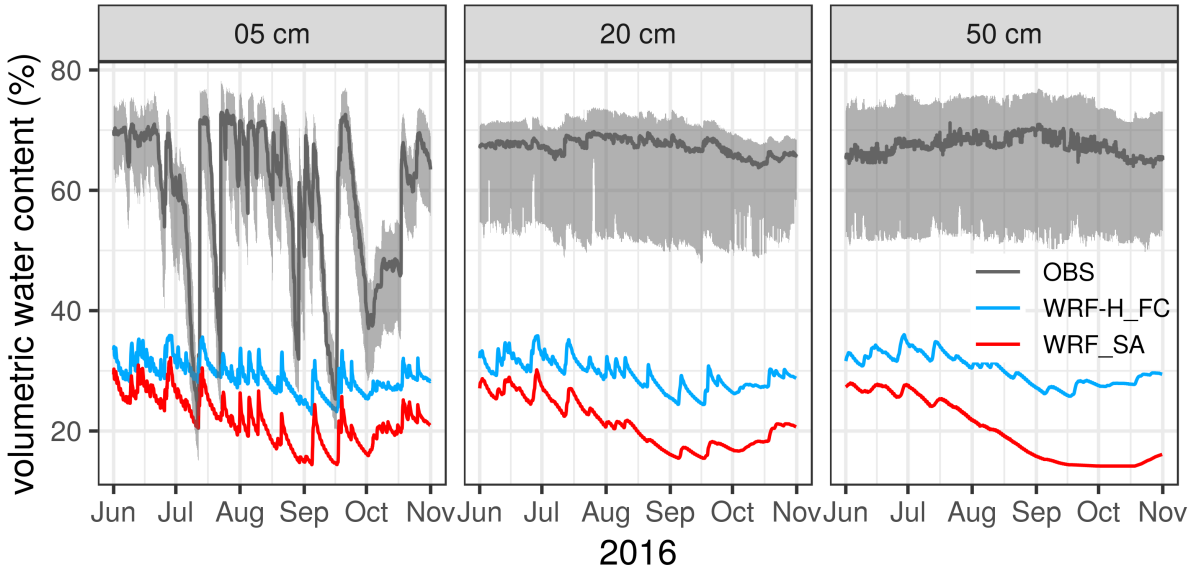

Figure 11. Volumetric water content at DE-Fen site for three different depths. Gray ribbon: range between first and third quartile of SoilNet observations; gray line: median of SoilNet observations

centiles of a wireless soil moisture sensor network that consists of 55 profiles with measurements at 5, 20, and 50 cm depth (SoilNet, further details are available in Kiese et al., 2018; Fersch et al., 2018). Obviously, simulations and observation show a considerable offset and also the temporal variations are much smoother for the model. The discrepancies in the soil moisture time-series are largely attributable to the difference in saturation water content. The LSM assumes loam for the DE-Fen site

(and almost for the entire area of domain 3) with a maximum volumetric soil moisture content of 44 % whereas in reality the region consists of sandy to silty loams and also peaty areas where the maximum soil moisture ranges between 50-80 %. With WRF-H_FC the soil moisture values are about 8-10 % higher than with WRF_SA. Also the decline differs between the two with WRF_SA leading to a much dryer scenario for the summer months. That is where WRF-H_FC-simulated latent heat flux and evapotranspiration largely outperform. Altogether, for DE-Fen, the decline predicted by WRF-H_FC seems more realistic with respect to the observations. This is also confirmed by the mostly improved statistical measures (Tab. 8). The representation of soil moisture in LSMs is a general challenge. Soil parameters and water contents are often tuned to unrealistic values for the sake of obtaining a good matching of the surface exchange fluxes with observations (Koster et al., 2009). The recent publications of global and continental high resolution soil hydraulic datasets (like, e.g., Hengl et al., 2017; Tóth et al., 2017) are helpful to improve and unify soil moisture representations in those models. However, these datasets, with their underlying water retention models (e.g., Van Genuchten, 1980), are not supported by the Noah LSMs and an implementation would be out of the scope of this study.

Table 8. Soil moisture performance measures (vol. %), DE-Fen, June-October 2016.

Depth	Model	r^2	ME	MAE
05 cm	WRF_SA	0.28	−37.22	37.23
	WRF-H_FC	0.29	−30.01	30.12
20 cm	WRF_SA	0.07	−46.03	46.03
	WRF-H_FC	0.11	−37.88	37.88
50 cm	WRF_SA	0.01	−47.41	47.41
	WRF-H_FC	0.02	−37.19	37.19

3.2 Boundary layer profiles during ScaleX campaign 2016

Figure 12 shows the spline interpolated vertical profiles of the performance measures for the WRF_SA and WRF-H_FC simulations and the HATPRO observations for the planetary boundary layer. The measurements represent hourly subsampled time-series from 2016-06-01 to 2016-07-31 for air temperature (Fig. 12 a, b, c) and absolute humidity (Fig. 12 d, e, f). For temperature, the differences between HATPRO and the models are generally much larger than the HATPRO accuracy. For humidity, the mean deviation lies within the accuracy of the measurement. It can be followed that both models overestimate temperature and probably have a tendency to underestimate absolute humidity. As compared to WRF_SA, WRF-H_FC shows reduced deviations for both variables.

The increase in correlation and the decrease of errors with height is only visible for temperature. The comparison between WRF_SA and WRF-H_FC reveals some modifications for the near surface region, where the fully coupled model gives a

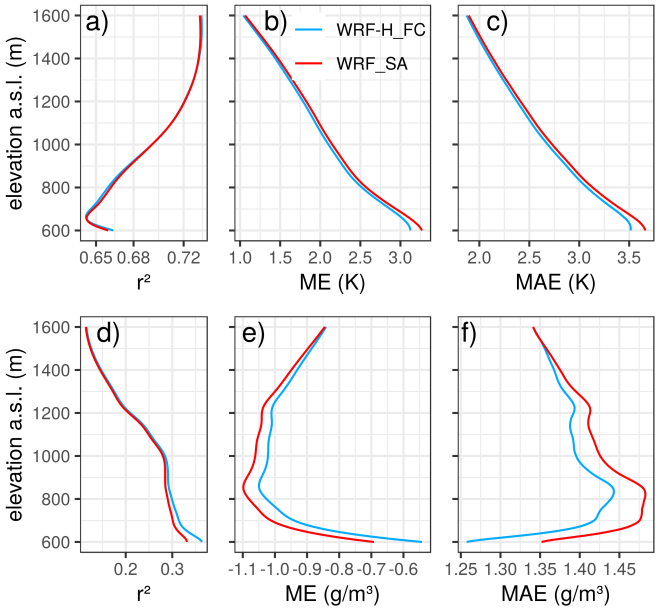

Figure 12. Performance evaluation of vertical atmospheric profiles of air temperature (upper panel) and absolute humidity (lower panel) for the DE-Fen site, 2016-06-01 to 2016-07-31. r^2= coefficent of determination, ME = mean error, MAE = mean absolute error.

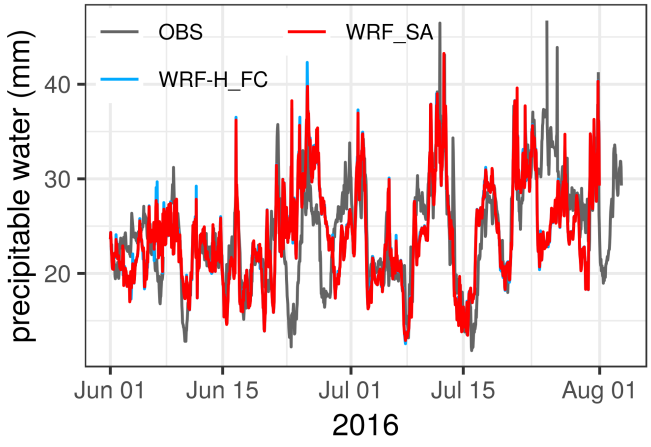

Figure 13. Simulated and observed precipitable water for the DE-Fen site, 2016-06-01 to 2016-07-31.

slightly improved skill. For absolute humidity, the lower parts of the profiles are in better agreement with the observations. With 0.15 to 0.35, the coefficients of determination are small as compared to 0.64 to 0.73 for temperature. The WRF-H_FC run outperforms WRF_SA especially for the lower 400 m of the profile. The time series of simulated and observed integrated water vapor for the DE-Fen site are shown in Figure 13. The temporal evolution is reasonably covered, with a few larger mismatches

at the end of June and the end of July. The inter-model differences are very small. Both simulations show nearly identical performance with r^2=0.42, ME=-0.21/-0.19 mm, and MAE=3.34/3.34 mm for WRF_SA and WRF-H_FC, respectively. The results for integrated water vapor and humidity profiles indicate that the coupling mainly affects the atmospheric boundary layer as differences in the correlation and errors between both simulations are restricted to the lower heights. Moreover, the domain area seems too small for internal moisture recycling and additional precipitation generation to take place. Most of the surplus in humidity is probably transported beyond the domain boundary. If the coupled simulation was extended to a larger area, e.g., Europe, impact above the boundary layer would be expected, at least for weak synoptic forcing periods (Arnault et al., 2018).

3.3 Water budgets

3.3.1 Analysis for sub-basin integrated water balances and discharge

Figure 14 visualizes the monthly water budgets for the six different subcatchments for WRF_SA, WRF-H_FC, and WRF-H_SA (stand alone WRF-Hydro) simulations, according to the water balance equation

$$P = E + R_{SF} + R_{UG} + \Delta S_{soil}$$

with precipitation P, evapotranspiration E, surface and subsurface runoff R_{SF} and R_{UG}, and soil storage variation ΔS_{soil}. Deviations in subcatchment aggregated precipitation are small for WRF_SA and WRF-H_FC. P in WRF-H_SA, that originates from the gridded observation product (RADOLAN), differs for most of the months and regions. WRF_SA and WRF-H_FC overestimate P for the mountainous regions (Am-OAG and Am-PEI) for May and June. For the other subcatchments June to August are mainly underestimated. September and October are well resembled for (Rt-RST). For the others, October sums are overestimated. For evapotranspiration E an increasing tendency can be seen from WRF_SA via WRF-H_FC to WRF-H_SA. Underestimated P in WRF_SA and WRF-H_FC results in decreased E. For the Am-OAG subcatchment, the overestimation in P is transferred to surface and subsurface runoff. This is likely because of the soil moisture in the mountain region is generally higher and E is rather energy limited and therefore cannot increase considerably. Moreover, slopes are steeper and soil thickness is reduced so that percolation takes place quickly. The variation in soil volumetric water content is irregular among models and for the different months. This indicates a non-linear feedback for the land–atmosphere interaction. Soil infiltration generally increases with the hydrologically enhanced models, however, the amounts for WRF-H_FC and WRF-H_SA vary according to P. Storage depletion (negative values) does not exhibit any tendency among the different models. Surface runoff (infiltration excess) with WRF_SA is 50% (for some of the subcatchments more than 100%) higher than with WRF-H_FC and WRF-H_SA. Conversely, groundwater recharge (soil drainage) increases for the hydrologically enhanced models. Again, differences between WRF-H_FC and WRF-H_SA are due to the individual precipitation amounts. On monthly scale, changes of the canopy water storage compensate (not shown). The water budget residuals, caused by the subgrid aggregation and disaggregation and by other numerical artifacts, can reach up to 31 mm for WRF-H_FC at Rt-RST in September but in the mean they are 5.6 mm. Altogether, the coupling with hydrology leads to increased infiltration, slightly increased E but almost no changes in P. A

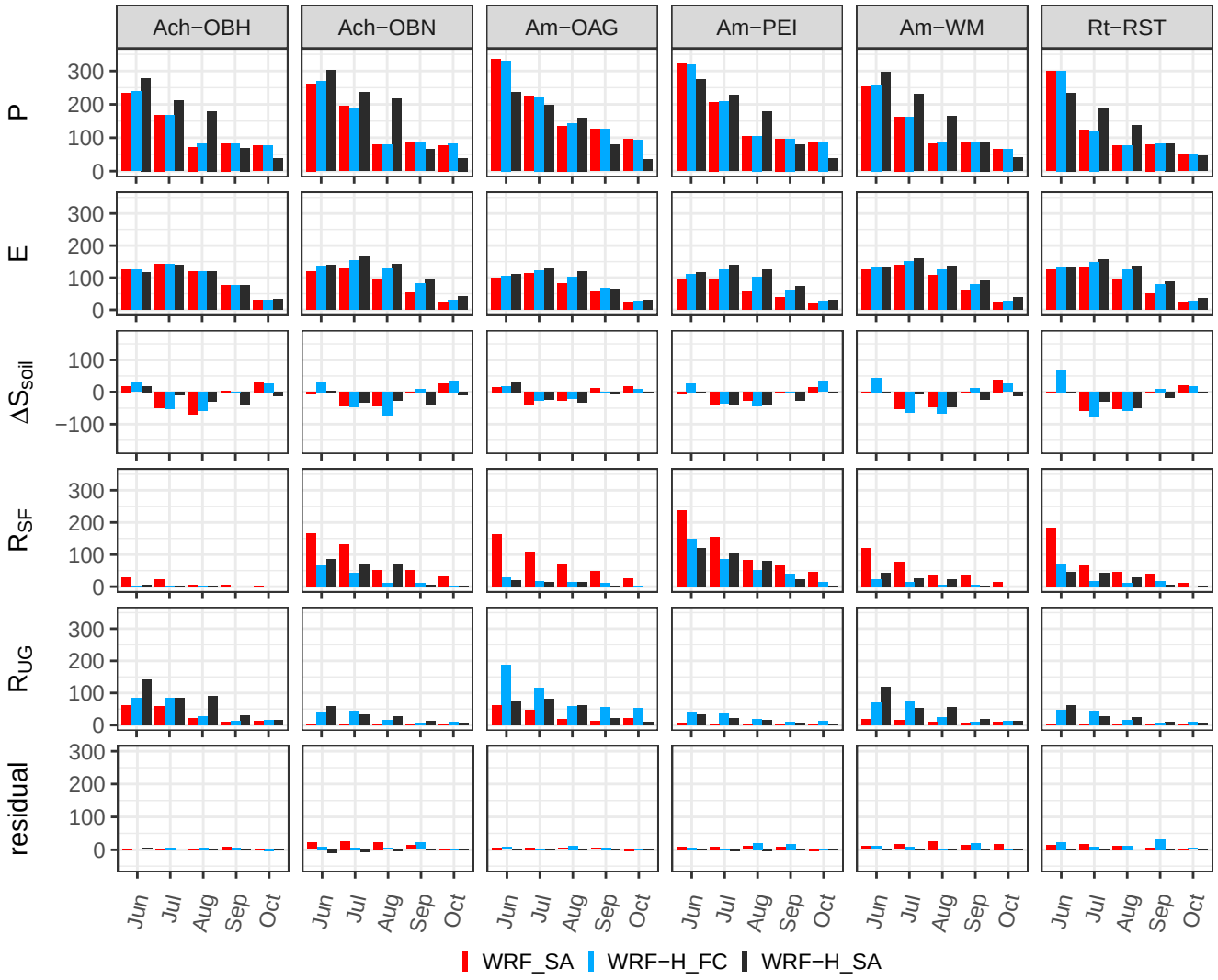

Figure 14. Monthly water budgets in mm for subcatchments Ach, Ammer (Am), and Rott (Rt). P = precipitation, E = evapotranspiration, ΔS_{soil} = soil water storage change, R_{sf} = surface runoff, R_{UG} = underground runoff (groundwater recharge). WRF_SA = standalone WRF model, WRF-H_FC = fully-coupled WRF-Hydro model, WRF-H_SA = observation driven WRF-Hydro standalone model.

reason for this could be that the distance of the displacement between precipitation generation and falling locations is generally much larger than the one covered by the domain boundaries in this study (Arnault et al., 2016a; Wei et al., 2015).

Table 9 lists the performance measures for the discharge simulated with the fully coupled model (hydrographs available in the supplementary material Fig. S3) with baseflow shifted simulations (compare Tab. 3, WRF-H_SA calibration) denoted by *sh*. Nash-Sutcliffe efficiencies (NSE) are poor for all stations, Kling-Gupta efficiencies (KGE', Kling et al., 2012) lie between 0.07 and 0.39 with a general performance gain for the baseflow shifted time-series. For all subcatchments except Am-OAG, negative MEs are observed, with values ranging from -22.6 to 2.98 mm mon^{-1}. For the three Ammer gauges (OAG, PEI, WM), adding the baseflow shifting leads to an improved baseline of the hydrographs and volumetric efficiency but also to considerable overestimation of the cumulated sums. If the non-shifted MEs for Q are compared with those of P it turns out that for some of the subcatchments (Am-WM, Ach-OBH, Rt-RST), the deviations are of similar amount as precipitation bias. The poor performance of the fully-coupled simulation to predict hourly discharge can be clearly mapped to the model's difficulty in reproducing the timing and positioning of precipitation.

Table 9. Discharge performance measures for subcatchment gauges modeled with the fully coupled WRF-H_FC model, June to October 2016. The gauges marked with *sh* denote the shifted discharge simulation results as applied in the WRF-H_SA calibration (Tab. 3). Mean errors (ME) for Q (discharge) and (P) are given in mm mon^{-1}.

Gauge	NSE	VE	KGE'	ME (Q)	ME (P)
Am-OAG	-1.61	0.44	0.14	2.98	17.20
Am-OAG_sh	-2.11	0.37	0.17	22.32	17.20
Am-PEI	-0.66	0.42	0.12	-16.62	1.70
Am-PEI_sh	-0.63	0.53	0.39	12.34	1.70
Am-WM	-0.56	0.46	0.25	-16.23	-13.90
Am-WM_sh	-0.55	0.52	0.35	12.43	-13.90
Ach-OBH	-0.22	0.44	0.31	-14.03	-12.10
Ach-OBN	-0.02	0.26	0.07	-22.60	-13.20
Rt-RST	-0.56	0.05	0.20	-1.55	-4.73

3.3.2 Spatial variations of simulated soil moisture patterns

Figure 15 shows the time-series of root mean square deviations (RMSD) of the spatial variograms for the WRF_SA and WRF-H_FC simulations subdivided by the four soil layers in the model. The variograms were computed using ten equidistant lags of 1 km from 1 to 10 km. The RMSDs were computed for the six different subcatchments and also for the Ach, Ammer and Rott catchments, and for the full domain. For the calculation, the subregions were masked so that the adjacent areas and lakes did not impact the results. The analysis reveals that the structural differences between the two models have their

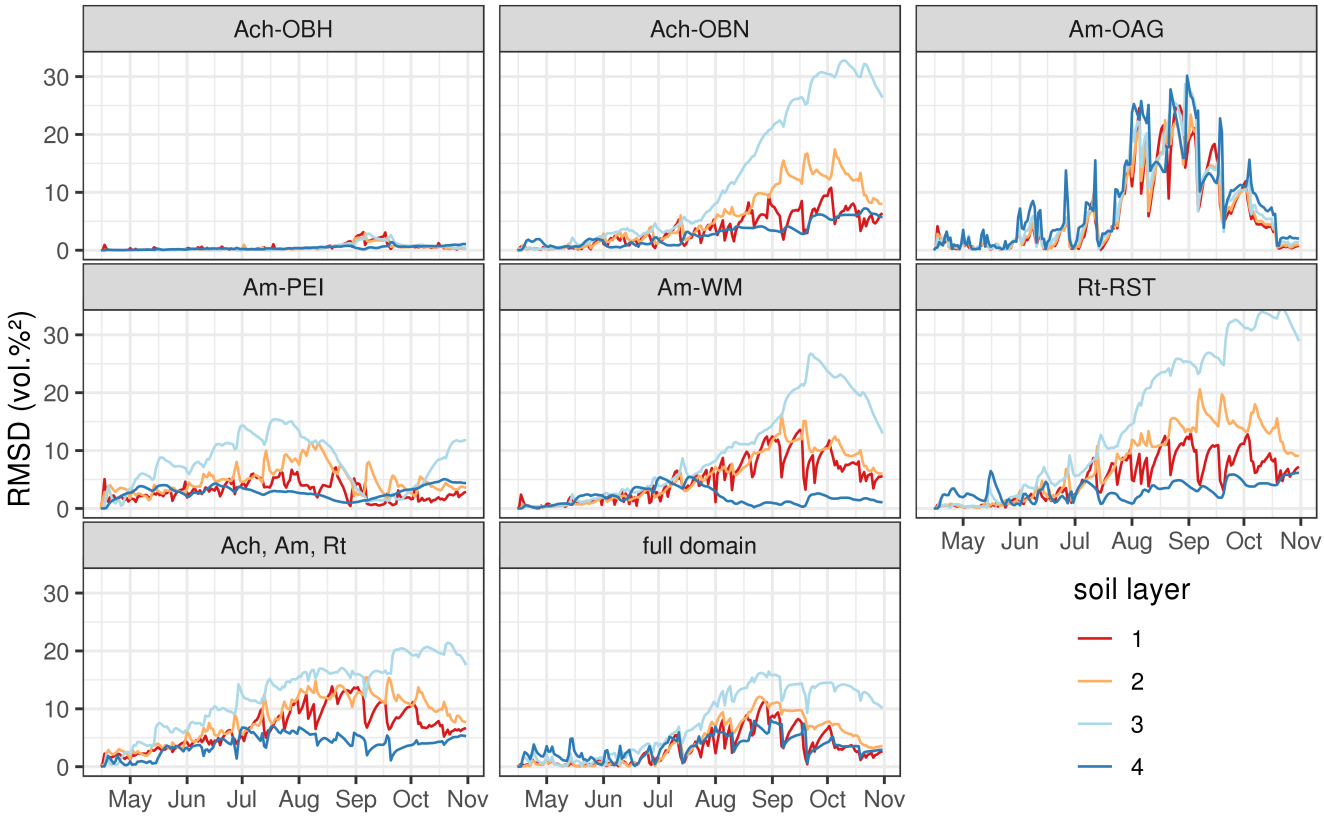

Figure 15. Time-series of root mean square deviations between empirical variograms of WRF_SA and WRF-H_FC for soil moisture layers 1 to 4, based on 10 equidistant lags (1 to 10 km).

maximum in late summer and fall. Surprisingly, layer three gives the strongest variations in spatial patterns. The changes for layers one and two are not so pronounced. The cause for this might be that the thinner top layers are strongly influenced by precipitation and infiltration processes and thus the spatial patterns are shaped accordingly. Due to its larger thickness, layer three reacts more sluggish. Thus deviations are more persistent. Layer four is even thicker but shows only slight variations over time. It is likely that the free drainage boundary condition has a regulatory effect here. Also the withdrawal of water for plant transpiration is partly reduced, as only forest landcover classes have roots in this layer per definition. Am-OAG depicts a special case as soil layers are considerably thinner for the steep mountain slopes (see 2.4.4). Thus, response times are short and routing of infiltration excess water is quickly propagated through all layers. The variability for the united Ach, Ammer and Rott catchments is less pronounced than seen for the smaller entities, but still the maximums are from late summer to fall, and layer three is affected most.

4 Summary, conclusive remarks and perspectives

The calibration of water related land surface parameters is hardly used for local area and regional climate model applications. The incorporation of water budgets in the model optimization provides an additional means to evaluate with independent observations. Such a concept requires a coupled atmospheric–hydrological approach that relates the land surface to planetary boundary layer exchange of energy and water with the spatial redistribution processes of water thus enabling the closure of the regional water balance and complex feedback processes at the land–atmosphere boundary.

This study examines the skills of a classic and a hydrologically-enhanced–fully-coupled setup of the Weather Research and Forecasting model (WRF / WRF-Hydro) to reproduce the land–atmosphere exchange of energy and water as well as the water budgets of the Ammer and Rott watersheds in southern Germany, for a six month period of 2016. The evaluation is based on comprehensive observations available by the TERENO Pre-Alpine environmental observatory and further third-party data suppliers such as the Bavarian Environmental Agency.

A standalone version of the WRF-Hydro model (without the atmospheric part), driven by WRF-simulated meteorological variables and observed precipitation (DWD-RADOLAN) was calibrated for six different subcatchments and the resulting parameters were subsequently used for a standalone WRF and a fully-coupled WRF-Hydro simulation both being identical with respect to initialization, parameters, forcing and binary code.

The calibration of the standalone WRF-Hydro model (WRF-H_SA) based on observed precipitation yielded reasonable results in terms of Nash-Sutcliffe and volumetric efficiencies. For the Ammer subcatchments (OAG, PEI, WM), due to long-term hydrogeological storage processes that cannot be reproduced by WRF-Hydro, it was required to correct the negative biases of the baseflow. The volumetric efficiency measure was an important indicator for further optimizing the parameters when Nash-Sutcliffe and Kling-Gupta efficiencies already converged.

For the validation period (2016-05-01–2016-10-31), the skill measures could largely outperform those of the calibration period. However, some structural deficiencies like the underestimated baseflow in the mountainous parts or negative bias remain. It is concluded that some of the processes in the catchment cannot be depicted because of the model physics or the lumped parameter estimation approach. Altogether, the subcatchment by subcatchment calibration is a very time-consuming effort, even if done in a semi-parallel way which is to our opinion not applicable for larger study regions such as on the national or continental level. A solution might be to switch to a land-characteristics-based universal method or to use a multi-scale parameter regionalization method as, for example, described in Mizukami et al. (2017).

For the fully-coupled WRF-Hydro run (WRF-H_FC), to obtain commensurable quantities for the evaluation, instead of calling the hydrological subgrid functions at every WRF model timestep (4 sec), an hourly timestep identical to that of the WRF-H_SA model was required. The strategy was selected to avoid numerical truncation effects in the overland routing routines that happen when timesteps are in the order of a few seconds and spatial resolution is about 100 m or below (see also section 2.4.4). The most prominent impact of the enabled lateral routing, on WRF-H_FCversus WRF_SAis a general increase in soil moisture values due to lateral water transport at the land surface which in turn leads to increased evapotranspiration for the summer months. Compared to the observations, the coupled simulation performs better for most of the months and this

finding holds also for the fluxes of sensible and ground heat. Solely for the mountainous site (DE-Gwg) both models show almost identical results which we attribute to generally higher soil moisture in that region and also to the reduced soil water storage capacity that comes with the decreased layer thicknesses defined for the slopy regions. In addition to the fluxes, also the near-surface states for air temperature and mixing ratio are better met with the fully-coupled model. The comparison with observed boundary layer temperature and humidity profiles at the DE-Fen site gives also a higher rank for the WRF-H_FC simulation. WRF_SA yields much dryer soil moisture conditions, and in particular from July to September it cannot maintain the evapotranspiration that is seen from the observations. However, when compared with the SoilNet observations at the DE-Fen TERENO Pre-Alpine site, a considerable mismatch remains ascribable to the discrepancy of the soil maps used in the model and real world conditions.

The STATSGO/FAO data assumes silty loam for almost the entire area of domain 3. This does not even rudimentarily reflect the complexity of the northern limestone Alps and foothills (compare Hofmann et al., 2009, for the Halbammer subcatchment of the Ammer), in particular with the high resolutions of 1 km and 100 m of the atmospheric and hydrological model sections. The incorporation of high-resolution continental or global soil maps like, e.g., recently made available by Hengl et al. (2017) and Tóth et al. (2017) could lead to further improvement here. The inter-model deviations for precipitation are marginal. The increased latent heat flux of the fully-coupled model does not strongly impact the precipitation generation for the study domain. It is rather assumed that the surplus in atmospheric moisture is actually being transported beyond the lateral boundaries. Thus, to see an impact of the model coupling on precipitation patterns, the domain size should be sufficiently enlarged. It should be noted that, however, the impact of the lateral water transport modeling on average precipitation amounts may be much less important than, e.g., the selection of the PBL scheme (Arnault et al., 2018). The findings are also corroborated by water vapor tagging studies that conclude that regional precipitation recycling is largest during weak synoptic forcing and that variations on larger scales are rather small (Wei et al., 2015; Arnault et al., 2019). Nevertheless, for larger regions the impact may be considerable. Short- and longwave radiation do not change much with the different model configurations. As for precipitation, the land–atmosphere feedback of moisture by LSM–hydrological coupling has no noticeable impact on the cloud generation processes, thus leaving the biases unaffected.

The analysis of the subcatchment water budgets reveals a clear connection between the biases of precipitation, soil infiltration, evapotranspiration, and discharge. The other terms of the water balance equation, soil water storage variation and percolation, do not show distinct trends between standard and coupled simulations which could be related to the inter-subcatchment variations of the infiltration and percolation parameters.

Using the calibrated parameters of WRF-H_SA in WRF-H_FC is required from a physical perspective. If deviation patterns reoccurred, a recalibration of WRF-H_FC could lead to improved discharge simulations, however at the risk of deteriorating the other water budgets.

In contrast to the uncoupled WRF model, fully-coupled simulations with an observation-calibrated hydrologically enhanced LSM show partly improved skill for land–atmosphere exchange variables, although the physical realism of the hydrological extension as well as for the spatial patterns of static data is still limited. Additional efforts are required to increase this physical realism, which should further improve the skill of the overall system. Nevertheless, including hydrological processes provides

an additional way to calibrate and evaluate the simulations by taking also the regional water balance and budgets into account – provided that comprehensive observations are available.

The evaluation included the standalone WRF (WRF_SA) and the fully-coupled (WRF-H_FC) models that share identical parameter sets, initial and boundary conditions for a commensurable set of simulations. It is of course possible that other parameter combinations for WRF_SA could lead to likewise or even better performance as with WRF-H_FC. E.g., the dryer soils could be alleviated by increasing the value of the infiltration parameter REFKDT which in turn would increase evapotranspiration and decrease sensible heat. WRF_SA could be parameterized with domain wide setting with respect to point observations (e.g., latent/heat fluxes or soil moisture measurements), which however do not guarantee the same accuracy in the remaining area and the effects of the intermediate surface water storage and lateral flow that are particularly important for strong precipitation events would not be considered. In contrast, the hydrological model allows to take into account representative features of a wide (catchment) area, summarized in the discharge variable.

The combined approach offers potential to improve also future Earth System Modeling, like also pointed out by Clark et al. (2015). To experience the full momentousness of coupled atmospheric–hydrological modeling future studies should be extended to larger regions to cover the scales of atmospheric recycling processes. Also the descriptions of the hydrological processes in the models should be further refined as computational capabilities increase and with more and more detailed data products becoming available.

Code and data availability

The source code of the extended coupled WRF-3.7.1/WRF-Hydro-3.0 model, as used in this study, is available at Fersch (2019c, https://doi.org/10.5281/zenodo.3405780). The model configuration files (WRF and WRF-Hydro namelists) can be obtained from Fersch (2019a, https://doi.org/10.5281/zenodo.3407166). The TERENO-preApline, and the ScaleX campaign datasets used in this study are available for download at Fersch (2019b, https://doi.org/10.5281/zenodo.3406970). The discharge observations used for model calibration published by the Bavarian Environmental Agency at https://www.gkd.bayern.de. The RADOLAN data of the German Weather Service (DWD) are available at https://opendata.dwd.de. The ASTER global digital elevation model can be obtained from https://doi.org/10.5067/ASTER/ASTGTM.003. ASTER GDEM is a product of Japan's Ministry of Economy, Trade, and Industry (METI) and NASA.

Author contributions. BF, AS, and HK developed the methodology for the study. AS conducted the initial WRF simulations and the PEST calibration of WRF-H_SA. BF implemented the changes to the WRF-Hydro model code and performed the WRF-H_SA and WRF-H_FC simulations. BA, MM, KS, and IV conducted and processed the HATPRO, eddy-covariance, lysimeter, and soil moisture observations, respectively. BF and AS carried out the investigation and prepared the original draft with contributions from all authors. HK supervised the project.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was funded by the Helmholtz Association Initiatives REKLIM, and TERENO, the Karlsruhe Institute of Technology, the University of Calabria, the German Academic Exchange Service (DAAD) and the German Research Foundation (DFG Research Unit *Cosmic Sense*, FOR 2694). Additional funding was obtained from the German Research Foundations (DFG ATMOWATER, KU 2090/10, DFG IMAP II KU2090/7-2). We are grateful to David Gochis and Wei Yu from National Center for Atmospheric Research, Boulder, CO for their valuable suggestions and their assistance with the WRF-Hydro model. Moreover, we acknowledge the developers of the open source tools used for this study: GNU R (tidyverse, ggplot, sp, rgdal, influxdbr, stargazer, hydroGOF, hydroPSO, parallel), WRF, inkscape, TEX, Qgis, influxdb, PEST. The WRF-Hydro modeling system was developed at the National Center for Atmospheric Research (NCAR) through grants from the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA). NCAR is sponsored by the United States National Science Foundation.

References

Arnault, J., Knoche, R., Wei, J., and Kunstmann, H.: Evaporation tagging and atmospheric water budget analysis with WRF: A regional precipitation recycling study for West Africa, Water Resour. Res., 52, 1544–1567, doi:10.1002/2015wr017704, 2016a.

Arnault, J., Wagner, S., Rummler, T., Fersch, B., Bliefernicht, J., Andresen, S., and Kunstmann, H.: Role of Runoff-Infiltration Partitioning and Resolved Overland Flow on Land-Atmosphere Feedbacks: A Case Study with the WRF-Hydro Coupled Modeling System for West Africa, J. Hydrometeor, 17, 1489–1516, doi:10.1175/JHM-D-15-0089.1, 2016b.

Arnault, J., Rummler, T., Baur, F., Lerch, S., Wagner, S., Fersch, B., Zhang, Z., Kerandi, N., Keil, C., and Kunstmann, H.: Precipitation Sensitivity to the Uncertainty of Terrestrial Water Flow in WRF-Hydro: An Ensemble Analysis for Central Europe, J. Hydrometeor., 19, 1007–1025, doi:10.1175/jhm-d-17-0042.1, https://doi.org/10.1175/JHM-D-17-0042.1, 2018.

Arnault, J., Wei, J., Rummler, T., Fersch, B., Zhang, Z., Jung, G., Wagner, S., and Kunstmann, H.: A joint soil-vegetation-atmospheric water tagging procedure with WRF-Hydro: Implementation and application to the case of precipitation partitioning in the upper Danube river basin, Water Resour. Res., 0, doi:10.1029/2019wr024780, 2019.

Barlage, M., Tewari, M., Chen, F., Miguez-Macho, G., Yang, Z.-L., and Niu, G.-Y.: The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP, Climatic Change, 129, 485–498, doi:10.1007/s10584-014-1308-8, 2015.

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., and Gerlach, N.: Projekt RADOLAN – Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer), Deutscher Wetterdienst, Hydrometeorologie, 5, 2004.

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrology and Earth System Sciences, 5, 1–12, doi:10.5194/hess-5-1-2001, 2001.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279–298, doi:10.1002/hyp.3360060305, 1992.

Beven, K. and Binley, A.: GLUE: 20years on, Hydrological Processes, 28, 5897–5918, doi:10.1002/hyp.10082, 2014.

Beven, K. J., Kirkby, M. J., Schofield, N., and Tagg, A. F.: Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments, Journal of Hydrology, 69, 119–143, doi:10.1016/0022-1694(84)90159-8, 1984.

Blöschl, G., Blaschke, A. P., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, A., Haas, P., Hogan, P., Kazemi Amiri, A., Oismüller, M., Parajka, J., Silasari, R., Stadler, P., Strauss, P., Vreugdenhil, M., Wagner, W., and Zessner, M.: The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory, Hydrol. Earth Syst. Sci., 20, 227–255, doi:10.5194/hess-20-227-2016, 2016.

Bogena, H. R., Huisman, J. A., Schilling, B., Weuthen, A., and Vereecken, H.: Effective calibration of low-cost soil water content sensors, Sensors, 17, 208, doi:10.3390/s17010208, 2017.

Büttner, G.: CORINE Land Cover and Land Cover Change Products, pp. 55–74, Springer Netherlands, Dordrecht, doi:10.1007/978-94-007-7969-3_5, 2014.

Butts, M., Drews, M., Larsen, M. A., Lerer, S., Rasmussen, S. H., Grooss, J., Overgaard, J., Refsgaard, J. C., Christensen, O. B., and Christensen, J. H.: Embedding complex hydrology in the regional climate system – Dynamic coupling across different modelling domains, Advances in Water Resources, 74, 166–184, doi:10.1016/j.advwatres.2014.09.004, 2014.

Cervarolo, G., Mendicino, G., and Senatore, A.: A coupled ecohydrological–three-dimensional unsaturated flow model describing energy, H2O and CO2 fluxes, Ecohydrology, 3, 205–225, doi:10.1002/eco.111, 2010.

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C., Swenson, S. C., and Zeng, X.: Improving the representation of hydrologic processes in Earth System Models, Water Resour. Res., 51, 5929–5956, doi:10.1002/2015wr017096, 2015.

Cohen, S., Praskievicz, S., and Maidment, D. R.: Featured Collection Introduction: National Water Model, J Am Water Resour Assoc, 54, 767–769, doi:10.1111/1752-1688.12664, 2018.

Corbari, C. and Mancini, M.: Calibration and Validation of a Distributed Energy–Water Balance Model Using Satellite Data of Land Surface Temperature and Ground Discharge Measurements, Journal of Hydrometeorology, 15, 376–392, doi:10.1175/JHM-D-12-0173.1, 2014.

Crewell, S. and Löhnert, U.: Accuracy of boundary layer temperature profiles retrieved with multifrequency multiangle microwave radiometry, IEEE Transactions on Geoscience and Remote Sensing, 45, 2195–2201, doi:10.1109/tgrs.2006.888434, 2007.

Criss, R. E. and Winston, W. E.: Do Nash values have value? Discussion and alternate proposals, Hydrol. Process., 22, 2723–2725, doi:10.1002/hyp.7072, 2008.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., KÃ¶hler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Dickinson, R. E., Shaikh, M., Bryant, R., and Graumlich, L.: Interactive Canopies for a Climate Model, J. Climate, 11, 2823–2836, doi:10.1175/1520-0442(1998)011<2823:icfacm>2.0.co;2, 1998.

Doherty, J.: PEST: a unique computer program for model-independent parameter optimisation, Water Down Under 94: Groundwater/Surface Hydrology Common Interest Papers; Preprints of Papers, p. 551, 1994.

Fersch, B.: Model configuration for the coupled WRF/WRF-Hydro model used for the HESS manuscript "High- resolution fully-coupled atmospheric–hydrological modeling: a cross-compartment regional water and energy cycle evaluation", doi:10.5281/zenodo.3407166, 2019a.

Fersch, B.: TERENO-preAlpine observatory and ScaleX 2016 campaign data set associated with HESS paper "High-resolution fully-coupled atmospheric–hydrological modeling: a cross-compartment regional water and energy cycle evaluation", doi:10.5281/zenodo.3406970, 2019b.

Fersch, B.: WRF-Hydro 3.0 / WRF 3.7.1 regional hydrometeorological model, extended by Karlsruhe Institute of Technology, Campus Alpin, doi:10.5281/zenodo.3405781, 2019c.

Fersch, B., Jagdhuber, T., Schrön, M., Völksch, I., and Jäger, M.: Synergies for Soil Moisture Retrieval Across Scales From Airborne Polarimetric SAR, Cosmic Ray Neutron Roving, and an In Situ Sensor Network, Water Resour. Res., 54, 9364–9383, doi:10.1029/2018wr023337, 2018.

Fetzer, K., Grottenthaler, W., Hofmann, B., Jerz, H., Rückert, G., Schmidt, F., and Wittmann, O.: Erläuterungen zur Standortkundlichen Bodenkarte von Bayern 1: 50.000 München–Augsburg und Umgebung; Hrsg.: Bayer. Geol. L, 1986.

Foken, T. and Wichura, B.: Tools for quality assessment of surface-based flux measurements, Agricultural and Forest Meteorology, 78, 83–105, doi:10.1016/0168-1923(95)02248-1, 1996.

Frank, H.: Glazial übertiefte Täler im Bereich des Isar-Loisach-Gletschers: Neue Erkenntnisse über Aufbau und Mächtigkeit des Quartärs in den alpinen Tälern, im Gebiet des Murnauer Schotters und im Tölzer Lobus (erste Mitteilung), E&G – Quaternary Science Journal, 29, doi:10.23689/fidgeo-1329, 1979.

Fu, J., Gasche, R., Wang, N., Lu, H., Butterbach-Bahl, K., and Kiese, R.: Impacts of climate and management on water balance and nitrogen leaching from montane grassland soils of S-Germany, Environmental Pollution, 229, 119–131, doi:10.1016/j.envpol.2017.05.071, 2017.

García-Díez, M., Fernández, J., and Vautard, R.: An RCM multi-physics ensemble over Europe: multi-variable evaluation to avoid error compensation, Climate Dynamics, 45, 3141–3156, doi:10.1007/s00382-015-2529-x, 2015.

Givati, A., Gochis, D., Rummler, T., and Kunstmann, H.: Comparing One-Way and Two-Way Coupled Hydrometeorological Forecasting Systems for Flood Forecasting in the Mediterranean Region, Hydrology, 3, 19, doi:10.3390/hydrology3020019, 2016.

Gochis, D. J. and Chen, F.: Hydrological Enhancements to the Community Noah Land Surface Model: Technical Description, Tech. Rep. NCAR/TN-454+STR, National Center for Atmospheric Research, doi:10.5065/D60P0X00, 2003.

Gochis, D. J., Yu, W., and Yates, D. N.: The WRF-Hydro Model Technical Description and User's Guide, Version 3.0. NCAR Technical Document, 120 pp., Tech. rep., National Center for Atmospheric Research, http://www.ral.ucar.edu/projects/wrf_hydro/, 2016.

Graeff, T., Zehe, E., Blume, T., Francke, T., and Schröder, B.: Predicting event response in a nested catchment with generalized linear models and a distributed watershed model, Hydrological Processes, 26, 3749–3769, doi:10.1002/hyp.8463, 2012.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLOS ONE, 12, 1–40, doi:10.1371/journal.pone.0169748, 2017.

Hofmann, M., Engelhardt, S., Huwe, B., and Stumpp, C.: Regionalizing soil properties in a catchment of the Bavarian Alps, European Journal of Forest Research, 128, 597–608, doi:10.1007/s10342-008-0242-6, 2009.

Hornberger, G. and Spear, R.: Approach to the preliminary analysis of environmental systems, J. Environ. Manage.; (United States), 12:1, 1981.

Hundecha, Y. and Bárdossy, A.: Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, Journal of Hydrology, 292, 281 – 295, doi:https://doi.org/10.1016/j.jhydrol.2004.01.002, 2004.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, D13 103, doi:10.1029/2008JD009944, d13103, 2008.

Jensen, K. H. and Refsgaard, J. C.: HOBE: The Danish Hydrological Observatory, Vadose Zone Journal, 17, doi:10.2136/vzj2018.03.0059, 2018.

Julien, P. Y., Saghafian, B., and Ogden, F. L.: RASTER-BASED HYDROLOGIC MODELING OF SPATIALLY-VARIED SURFACE RUNOFF1, JAWRA Journal of the American Water Resources Association, 31, 523–536, doi:10.1111/j.1752-1688.1995.tb04039.x, 1995.

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, Journal of Applied Meteorology, 43, 170–181, doi:10.1175/1520-0450(2004)043<0170:tkcpau>2.0.co;2, 2004.

Katragkou, E., García-Díez, M., Vautard, R., Sobolowski, S., Zanis, P., Alexandri, G., Cardoso, R. M., Colette, A., Fernandez, J., Gobiet, A., Goergen, K., Karacostas, T., Knist, S., Mayer, S., Soares, P. M. M., Pytharoulis, I., Tegoulias, I., Tsikerdekis, A., and Jacob, D.: Regional climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble, Geosci. Model Dev., 8, 603–618, doi:10.5194/gmd-8-603-2015, 2015.

Kelleher, C., McGlynn, B., and Wagener, T.: Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding, Hydrology and Earth System Sciences, 21, 3325–3352, doi:10.5194/hess-21-3325-2017, 2017.

Kerandi, N., Arnault, J., Laux, P., Wagner, S., Kitheka, J., and Kunstmann, H.: Joint atmospheric-terrestrial water balances for East Africa: a WRF-Hydro case study for the upper Tana River basin, Theoretical and Applied Climatology, pp. 1–19, doi:10.1007/s00704-017-2050-8, 2017.

Kiese, R., Fersch, B., Bassler, C., Brosy, C., Butterbach-Bahlc, K., Chwala, C., Dannenmann, M., Fu, J., Gasche, R., Grote, R., Jahn, C., Klatt, J., Kunstmann, H., Mauder, M., Roediger, T., Smiatek, G., Soltani, M., Steinbrecher, R., Voelksch, I., Werhahn, J., Wolf, B., Zeeman, M., and Schmid, H.: The TERENO Pre-Alpine Observatory: Integrating Meteorological, Hydrological, and Biogeochemical Measurements and Modeling, Vadose Zone Journal, 17, doi:10.2136/vzj2018.03.0060, 2018.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424-425, 264–277, doi:10.1016/j.jhydrol.2012.01.011, 2012.

Kormann, R. and Meixner, F. X.: An Analytical Footprint Model For Non-Neutral Stratification, Boundary-Layer Meteorology, 99, 207–224, doi:10.1023/a:1018991015119, 2001.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., and Puma, M. J.: On the Nature of Soil Moisture in Land Surface Models, J. Climate, 22, 4322–4335, doi:10.1175/2009jcli2832.1, 2009.

Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., Goodrich, D., and Hazenberg, P.: Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments, J. Hydrometeor., 20, 691–714, doi:10.1175/jhm-d-18-0064.1, 2019.

Larsen, M. A. D., Rasmussen, S. H., Drews, M., Butts, M. B., Christensen, J. H., and Refsgaard, J. C.: Assessing the influence of groundwater and land surface scheme in the modelling of land surface-atmosphere feedbacks over the FIFE area in Kansas, USA, Environmental Earth Sciences, 75, 130, doi:10.1007/s12665-015-4919-0, 2016a.

Larsen, M. A. D., Refsgaard, J. C., Jensen, K. H., Butts, M. B., Stisen, S., and Mollerup, M.: Calibration of a distributed hydrology and land surface model using energy flux measurements, Agricultural and Forest Meteorology, 217, 74–88, doi:10.1016/j.agrformet.2015.11.012, 2016b.

Li, L., Gochis, D. J., Sobolowski, S., and Mesquita, M. D. S.: Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model, J. Geophys. Res. Atmos., 122, 4786–4807, 2017.

Löhnert, U. and Crewell, S.: Accuracy of cloud liquid water path from ground-based microwave radiometry 1. Dependency on cloud model statistics, Radio Science, 38, doi:10.1029/2002RS002654, 2003.

Löhnert, U., Turner, D. D., and Crewell, S.: Ground-Based Temperature and Humidity Profiling Using Spectral Infrared and Microwave Observations. Part I: Simulated Retrieval Performance in Clear-Sky Conditions, Journal of Applied Meteorology and Climatology, 48, 1017–1032, doi:10.1175/2008jamc2060.1, 2009.

Ludwig, R. and Mauser, W.: Modelling catchment hydrology within a GIS based SVAT-model framework, Hydrol. Earth Syst. Sci., 4, 239–249, doi:10.5194/hess-4-239-2000, 2000.

Marx, A.: Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, vol. 160 of *Mitteilungen / Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart*, Universität Stuttgart, doi:10.18419/opus-256, 2007.

Mauder, M. and Foken, T.: Eddy-Covariance Software TK3, doi:10.5281/zenodo.20349, 2015.

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, Agricultural and Forest Meteorology, 169, 122–135, doi:10.1016/j.agrformet.2012.09.006, 2013.

Maxwell, R. M. and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change, Nature Geoscience, 1, 665, doi:10.1038/ngeo315, 2008.

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resour. Res., 53, 8020–8040, doi:10.1002/2017wr020401, 2017.

Moore, C. J.: Frequency response corrections for eddy correlation systems, Boundary-Layer Meteorology, 37, 17–35, doi:10.1007/bf00122754, 1986.

Naabil, E., Lamptey, B. L., Arnault, J., Kunstmann, H., and Olufayo, A.: Water resources management using the WRF-Hydro modelling system: Case-study of the Tono dam in West Africa, Journal of Hydrology: Regional Studies, 12, 196–209, doi:10.1016/j.ejrh.2017.05.010, 2017.

Ning, L., Zhan, C., Luo, Y., Wang, Y., and Liu, L.: A review of fully coupled atmosphere-hydrology simulations, Journal of Geographical Sciences, 29, 465–479, doi:10.1007/s11442-019-1610-5, 2019.

Niu, G.-Y.: The Community NOAH Land-surface Model (LSM) with Multi-physics Options, User's Guide Public Release Version, 2, 2011.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res., 116, D12 109, doi:10.1029/2010JD015139, 2011.

Ogden, F.: CASC2D Reference Manual, Tech. rep., University of Connecticut, Storrs, CT, 1997.

Peters, A., Nehls, T., Schonsky, H., and Wessolek, G.: Separating precipitation and evapotranspiration from noise – a new filter routine for high-resolution lysimeter data, Hydrology and Earth System Sciences, 18, 1189–1198, doi:10.5194/hess-18-1189-2014, 2014.

Pospichal, B. and Crewell, S.: Boundary layer observations in West Africa using a novel microwave radiometer, Meteorologische Zeitschrift, 16, 513–523, doi:10.1127/0941-2948/2007/0228, 2007.

Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L.: Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resources Research, 52, 7779–7792, doi:10.1002/2016WR019430, 2016.

Rose, T., Crewell, S., Löhnert, U., and Simmer, C.: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, Atmospheric Research, 75, 183 – 200, doi:https://doi.org/10.1016/j.atmosres.2004.12.005, cLIWA-NET: Observation and Modelling of Liquid Water Clouds, 2005.

Rummler, T., Arnault, J., Gochis, D., and Kunstmann, H.: Role of Lateral Terrestrial Water Flow on the Regional Water Cycle in a Complex Terrain Region: Investigation With a Fully Coupled Model System, J. Geophys. Res. Atmos., 124, 507— 529, doi:10.1029/2018jd029004, 2018.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, W05 523, doi:10.1029/2008WR007327, 2010.

Schaake, J. C., Koren, V. I., Duan, Q.-Y., Mitchell, K., and Chen, F.: Simple water balance model for estimating runoff at different spatial and temporal scales, J. Geophys. Res., 101, 7461–7475, doi:10.1029/95JD02892, 1996.

Schotanus, P., Nieuwstadt, F., and De Bruin, H.: Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes, Boundary-Layer Meteorology, 26, 81–93, doi:10.1007/bf00164332, 1983.

Schulla, J. and Jasper, K.: Model description WaSiM-ETH, Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich, 2007.

Seiler, K.-P.: Glazial übertiefte Talabschnitte in den Bayerischen Alpen: Ergebnisse glazialgeologischer, hydrologischer und geophysikalischer Untersuchungen, E&G – Quaternary Science Journal, 29, doi:10.23689/fidgeo-915, 1979.

Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N., and Kunstmann, H.: Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales, J. Adv. Model. Earth Syst., 7, 1693–1715, doi:10.1002/2015MS000510, 2015.

Shrestha, P., Sulis, M., Masbou, M., Kollet, S., and Simmer, C.: A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow, Monthly Weather Review, 142, 3466–3483, doi:10.1175/MWR-D-14-00029.1, 2014.

Silver, M., Karnieli, A., Ginat, H., Meiri, E., and Fredj, E.: An innovative method for determining hydrological calibration parameters for the WRF-Hydro model in arid regions, Environmental Modelling & Software, 91, 47–69, doi:10.1016/j.envsoft.2017.01.010, 2017.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, Journal of Computational Physics, 227, 3465–3485, doi:10.1016/j.jcp.2007.01.037, 2008.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, M., Duda, K. G., Huang, X. Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 3, Tech. rep., National Center for Atmospheric Research, 2008.

Soltani, M., Laux, P., Mauder, M., and Kunstmann, H.: Inverse distributed modelling of streamflow and turbulent fluxes: A sensitivity and uncertainty analysis coupled with automatic optimization, Journal of Hydrology, 571, 856–872, doi:10.1016/j.jhydrol.2019.02.033, 2019.

Spath, F., Kremer, P., Wulfmeyer, V., Streck, T., and Behrendt, A.: The Land Atmosphere Feedback Observatory (LAFO): A novel sensor network to improve weather forecasting and climate models, in: AGU Fall Meeting Abstracts, vol. 2018, pp. A41D–01, 2018.

Stisen, S., McCabe, M. F., Refsgaard, J. C., Lerer, S., and Butts, M. B.: Model parameter analysis using remotely sensed pattern information in a multi-constraint framework, Journal of Hydrology, 409, 337–349, doi:10.1016/j.jhydrol.2011.08.030, 2011.

Sukoriansky, S., Galperin, B., and Perov, V.: 'Application of a New Spectral Theory of Stably Stratified Turbulence to the Atmospheric Boundary Layer over Sea Ice', Boundary-Layer Meteorology, 117, 231–257, doi:10.1007/s10546-004-6848-4, 2005.

Sulis, M., Keune, J., Shrestha, P., Simmer, C., and Kollet, S. J.: Quantifying the Impact of Subsurface-Land Surface Physical Processes on the Predictive Skill of Subseasonal Mesoscale Atmospheric Simulations, J. Geophys. Res. Atmos., 123, 9131–9151, doi:10.1029/2017jd028187, 2018.

Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, Monthly Weather Review, 136, 5095–5115, doi:10.1175/2008MWR2387.1, https://doi.org/10.1175/2008MWR2387.1, 2008.

Thyer, M., Beckers, J., Spittlehouse, D., Alila, Y., and Winkler, R.: Diagnosing a distributed hydrologic model for two high-elevation forested catchments based on detailed stand- and basin-scale data, Water Resources Research, 40, W01 103, doi:10.1029/2003WR002414, 2004.

Tóth, B., Weynants, M., Pásztor, L., and Hengl, T.: 3D soil hydraulic database of Europe at 250 m resolution, Hydrological Processes, 31, 2662–2666, doi:10.1002/hyp.11203, 2017.

Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils 1, Soil science society of America journal, 44, 892–898, 1980.

Van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and Srinivasan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models, Journal of hydrology, 324, 10–23, doi:10.1016/j.jhydrol.2005.09.008, 2006.

Wagner, S., Fersch, B., Yuan, F., Yu, Z., and Kunstmann, H.: Fully coupled atmospheric-hydrological modeling at regional and long-term scales: Development, application, and analysis of WRF-HMS, Water Resour. Res., 52, 3187–3211, doi:10.1002/2015WR018185, 2016.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Quarterly Journal of the Royal Meteorological Society, 106, 85–100, doi:10.1002/qj.49710644707, 1980.

Wei, J., Knoche, H. R., and Kunstmann, H.: Contribution of transpiration and evaporation to precipitation: An ET-Tagging study for the Poyang Lake region in Southeast China, J. Geophys. Res. Atmos., 120, 6845–6864, doi:10.1002/2014jd022975, 2015.

Wigmosta, M. S. and Lettenmaier, D. P.: A comparison of simplified methods for routing topographically driven subsurface flow, Water Resources Research, 35, 255–264, doi:10.1029/1998WR900017, 1999.

Wigmosta, M. S., Vail, L. W., and Lettenmaier, D. P.: A distributed hydrology-vegetation model for complex terrain, Water Resources Research, 30, 1665–1679, doi:10.1029/94WR00436, 1994.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic Anemometer Tilt Correction Algorithms, Boundary-Layer Meteorology, 99, 127–150, doi:10.1023/a:1018966204465, 2001.

Winterrath, T., Rosenow, W., and Weigl, E.: On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management, Weather Radar and Hydrology, IAHS Publ, 351, 323–329, 2012.

Wolf, B., Chwala, C., Fersch, B., Garvelmann, J., Junkermann, W., Zeeman, M. J., Angerer, A., Adler, B., Beck, C., Brosy, C., Brugger, P., Emeis, S., Dannenmann, M., De Roo, F., Diaz-Pines, E., Haas, E., Hagen, M., Hajnsek, I., Jacobeit, J., Jagdhuber, T., Kalthoff, N., Kiese, R., Kunstmann, H., Kosak, O., Krieg, R., Malchow, C., Mauder, M., Merz, R., Notarnicola, C., Philipp, A., Reif, W., Reineke, S., Rüdiger, T., Ruehr, N., Schäfer, K., Schrön, M., Senatore, A., Shupe, H., Völksch, I., Wanninger, C., Zacharias, S., and Schmid, H. P.: The ScaleX campaign: scale-crossing land-surface and boundary layer processes in the TERENO-preAlpine observatory, Bull. Amer. Meteor. Soc., 98, 1217–1234, doi:10.1175/bams-d-15-00277.1, 2016.

Yucel, I., Onen, A., Yilmaz, K. K., and Gochis, D. J.: Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall, Journal of Hydrology, 523, 49–66, doi:10.1016/j.jhydrol.2015.01.042, 2015.

Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J. C., Papen, H., Priesack, E., Schmid, H. P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, Vadose Zone Journal, 10, 955–973, doi:10.2136/vzj2010.0139, 2011.