

RESEARCH ARTICLE

10.1002/2013WR014634

Special Section:

Patterns in
Soil-Vegetation-Atmosphere
Systems: Monitoring,
Modelling and Data
Assimilation

Key Points:

- Selecting different model options strongly influences accuracy of simulations
- The ensemble size can be reduced by constraining Noah-MP to different data types
- Considering dynamics of root growth results in more accurate simulations

Correspondence to:

S. Gayler,
sebastian.gayler@uni-tuebingen.de

Citation:

Gayler, S., T. Wöhling, M. Grzeschik, J. Ingwersen, H.-D. Witzmann, K. Warrach-Sagi, P. Högy, S. Attinger, T. Streck, and V. Wulfmeyer (2014), Incorporating dynamic root growth enhances the performance of Noah-MP at two contrasting winter wheat field sites, *Water Resour. Res.*, *50*, 1337–1356, doi:10.1002/2013WR014634.

Received 22 AUG 2013

Accepted 28 JAN 2014

Accepted article online 3 FEB 2014

Published online 18 FEB 2014

Incorporating dynamic root growth enhances the performance of Noah-MP at two contrasting winter wheat field sites

Sebastian Gayler¹, Thomas Wöhling^{1,2}, Matthias Grzeschik¹, Joachim Ingwersen³, Hans-Dieter Witzmann⁴, Kirsten Warrach-Sagi^{1,4}, Petra Högy⁵, Sabine Attinger^{1,6}, Thilo Streck^{1,3}, and Volker Wulfmeyer^{1,4}

¹Water and Earth System Science Competence Cluster (WESS), Tübingen, Germany, ²Lincoln Agritech Ltd., Ruakura Research Centre, Hamilton, New Zealand, ³Institute of Soil Science and Land Evaluation, Biogeophysics, University of Hohenheim, Stuttgart, Germany, ⁴Institute of Physics and Meteorology, University of Hohenheim, Stuttgart, Germany, ⁵Institute of Landscape and Plant Ecology, University of Hohenheim, Stuttgart, Germany, ⁶Computational Hydrosystems, UFZ-Helmholtz Centre for Environmental Research, Leipzig, Germany

Abstract Interactions between the soil, the vegetation, and the atmospheric boundary layer require close attention when predicting water fluxes in the hydrogeosystem, agricultural systems, weather, and climate. However, land-surface schemes used in large-scale models continue to show deficiencies in consistently simulating fluxes of water and energy from the subsurface through vegetation layers to the atmosphere. In this study, the multiphysics version of the Noah land-surface model (Noah-MP) was used to identify the processes, which are most crucial for a simultaneous simulation of water and heat fluxes between land surface and the lower atmosphere. Comprehensive field data sets of latent and sensible heat fluxes, ground heat flux, soil moisture, and leaf area index from two contrasting field sites in South-West Germany are used to assess the accuracy of simulations. It is shown that an adequate representation of vegetation-related processes is the most important control for a consistent simulation of energy and water fluxes in the soil-plant-atmosphere system. In particular, using a newly implemented submodule to simulate root growth dynamics has enhanced the performance of Noah-MP. We conclude that further advances in the representation of leaf area dynamics and root/soil moisture interactions are the most promising starting points for improving the simulation of feedbacks between the subsoil, land surface and atmosphere in fully coupled hydrological and atmospheric models.

1. Introduction

A detailed understanding of the interactions between soil, vegetation, and the atmospheric boundary layer is a prerequisite for predicting the effects of land use and climate change on hydrological systems. To better understand the complex interplay of the involved processes, many numerical and physics-based land-surface models (LSMs) have been developed during the last decades [Sellers *et al.*, 1986; Chen and Dudhia, 2001; Kothavala *et al.*, 2005; Krinner *et al.*, 2005; Oleson *et al.*, 2008; Bonan *et al.*, 2011; Niu *et al.*, 2011]. LSMs are used in hydrological studies at catchment and river basin scale [Lohmann *et al.*, 1998; Maxwell and Miller, 2005; Grasset *et al.*, 2008; Rosero *et al.*, 2011; Wolf, 2011], and in regional and global-scale hydrological studies [Wood *et al.*, 1998; Yang *et al.*, 2011], climate studies [Koster *et al.*, 2006; Hohenegger *et al.*, 2009; Breuer *et al.*, 2012; Dirmeyer, 2013; Greve *et al.*, 2013], and climate impact studies [Aurbacher *et al.*, 2013; Challinor *et al.*, 2013]. In particular, analyzing feedback processes in the soil-plant-atmosphere continuum requires a consistent description of the fluxes of water, energy, and carbon within and between the different components of coupled models [Santanello *et al.*, 2011]. However, considerable deficiencies in land-surface models in simultaneously simulating soil moisture, water fluxes, heat, and their mutual relations in the soil-plant continuum have been identified [Dirmeyer *et al.*, 2006; Koster *et al.*, 2006; Warrach-Sagi *et al.*, 2008; Niu *et al.*, 2011; Gayler *et al.*, 2013]. Thus, further improvement of land-surface models remains a major challenge.

A typical way to improve LSMs is to replace existing representations of biogeophysical processes with more advanced ones. With respect to vegetation processes, more detailed process representations can be adopted from field-scale crop and forest models, which typically have a greater emphasis on seasonal dynamics of root and leaf development than land-surface models, which are designed primarily for

large-scale applications. In particular, root water uptake is often poorly represented in LSMs [Overgaard *et al.*, 2006]. Gayler *et al.* [2013] showed that simplifications in the representation of root growth and root water uptake in the Community Land Model CLM 3.5 result in poor simulations of the dynamic and vertical distribution of soil moisture. Shortcomings of land-surface models can also be ameliorated by introducing additional processes, if poor simulations can be attributed to the absence of these processes. The latter way was for example pursued by Noblet-Ducoudré *et al.* [2004], who coupled the agronomy model STICS to the soil-vegetation-transfer scheme ORCHIDEE. The purpose of the latter study was to investigate the influence of croplands on the European carbon and water budgets.

A multitude of process parameterization options exists for the different biogeophysical processes simulated in land-surface models. In many cases, there is no clear reason for preferring one process representation over another. Often it is questionable whether improving the representation of a single process will also improve overall performance of a LSM, because undesirable interactions of parameters in the new scheme with ones in existing schemes of the LSM may occur [Niu *et al.*, 2011]. Moreover, when used coupled to an atmospheric model, changes in a LSM may have considerable impacts on the coupling with the atmosphere via parameterizations of the atmospheric surface layer. Aside from ambiguous decision about the overall model structure during the process of model development, the error associated with alternative representations of the individual process is a major contributor to structural model uncertainty [Clark *et al.*, 2011a]. The level of this uncertainty can be quite large and even greater than the uncertainty associated with input and parameter uncertainty [Ajami *et al.*, 2007]. Consequently, to quantify uncertainty in simulations of biogeophysical processes, multimodel or multiphysics approaches are needed in addition to techniques based on parameter perturbation.

Working with multimodel ensembles has a long tradition in climate research. It is increasingly becoming a common practice also in studies of the impact of climate change on hydrology and agroecosystems [Challinor *et al.*, 2013]. However, multimodel and multiphysics ensembles are more accustomed to assessing uncertainties in the climate forcing than to quantifying structural uncertainty in simulation schemes for simulating biogeophysical processes [Hemming *et al.*, 2013; Höglind *et al.*, 2013]. Only a few multimodel studies have been conducted to investigate the structural uncertainty of LSMs [Yang *et al.*, 2011; Zhou *et al.*, 2012]. The necessity of multimodel studies is also in the focus of ecophysiological research [Rötter *et al.*, 2011; Rosenzweig *et al.*, 2013]. The importance of structural model uncertainty in the modeling of agroecosystems under future climate conditions was addressed within the scope of the Wheat Pilot Study of the Agricultural Model Intercomparison and Improvement Project (AgMIP). In a study with 27 crop models, it was shown that the uncertainty in climate change projections of wheat yields was greater due to variations among models than to variations in the forcing provided by different climate models [Asseng *et al.*, 2013]. The impact of different complexity in representing crop growth of five coupled soil-plant models and of the Community Land Model (CLM 3.5) on the simultaneous simulation of soil moisture, evapotranspiration, and leaf area index at plot scale was recently analyzed by Wöhling *et al.* [2013].

To enhance the realism of biophysical and hydrological processes in the community land-surface model Noah, this model was recently enhanced by the addition of a multioptions framework that allows the user to select between different submodules for individual processes [Niu *et al.*, 2011]. The new version of the model is called Noah-MP (multiphysics). However, vegetation dynamics are still poorly represented. In previous studies, we found that the simulation of root growth and activity enhances the performance of coupled land-surface models and soil-vegetation models [Gayler *et al.*, 2013; Wöhling *et al.*, 2013]. We therefore extended the multioptions framework of Noah-MP with a new option that allows for the simulation of root growth of crops. In this study, we evaluate the potential of the new option to improve the performance of a Noah-MP multiphysics ensemble to simulate water and energy fluxes across the land surface at two contrasting agricultural field sites in South-West Germany. For this, we selected a multitude of combinations of alternative representations of the most critical processes that affect land-surface fluxes and soil moisture dynamics. We evaluated Noah-MP in the 1-D stand alone mode, i.e., without coupling to an atmospheric model. Instead, the atmospheric forcing is provided by measured time series of the relevant atmospheric data. Testing the performance of LSMs at the plot scale and driven by experimental climate data is a common practice, because the data needed for model evaluation can be gathered with relatively high accuracy and in high-temporal resolution [Dirmeyer *et al.*, 2006; Schädler, 2007;

Mahecha et al., 2010; Ingwersen et al., 2011]. Simulation results are compared with eddy covariance (EC) measurements of turbulent fluxes of water and energy, measurements of ground heat flux, and time series of soil moisture at different depths. The aims of the study are (i) to assemble the most appropriate combinations of process parameterizations in the Noah-MP model to simultaneously match the different components of the water and energy cycle at the field sites under consideration, (ii) to test the performance with the new root growth option and to identify further processes the modeling of which is most promising for model improvement, (iii) to recognize inconsistencies in the model structure and how they are affected by the new root growth option, (iv) to analyze whether the results found at one site are transferable to another, and (v) to estimate uncertainties in model predictions originating from alternative sub-models available in Noah-MP.

2. Materials and Methods

2.1. Field Data

Plot-scale data on soil moisture, ground heat flux (GHF), latent heat flux (LHF), and sensible heat fluxes (SHF), and leaf area development were recorded during the 2009 vegetation period at two open and flat agricultural fields sites in South-West Germany. The two study sites in the Kraichgau (48.9°N and 8.7°E, 319 m a.s.l.) and Swabian Alb (48.5°N and 9.8°E, 690 m a.s.l.) differ in terms of both their soils and climatic conditions. The Kraichgau is a fertile hilly loess region, characterized by a mild climate with comparatively high temperatures and moderate precipitation (mean annual: temperature 9.3°C, precipitation 777 mm) and intensive agricultural land use with high crop yields. The Swabian Alb is a mountain plateau with extensive agricultural land use. Its climate is distinctly colder and wetter (mean annual: temperature 6.5°C, precipitation 962 mm). The two agricultural fields (EC1, Kraichgau, 15 ha, and EC6, Swabian Alb, 13 ha) presented here are managed and operated by local farmers. The soil at EC1 is a Stagnic Anthrosol [IUSS Working Group WRB, 2007] on a loess layer of several meters depth. Winter wheat (*Triticum aestivum* cv. Cubus) was sown on 7 November 2008 and harvested on 6 August 2009 at grain maturity. EC6 is characterized by a shallow and rendzic Leptosol [IUSS Working Group WRB, 2007] with a solum depth of 0.2–0.3 m. On 7 October 2008 winter wheat (cv. Hermann) was sown. Harvest took place at 24 August 2009. On each of the fields, five subplots of 4 m² were randomly selected in early spring 2009 and permanently marked to track crop performance (phenology and leaf area index (LAI)). LAI was measured on each subplot at the central square meter in biweekly intervals until grain maturity using a LAI-2000 Plant Canopy Analyzer (LI-COR Biosciences Inc., USA). Energy and water fluxes between canopy and atmosphere were measured with the EC technique. Both stations were equipped with the same set of instruments. The instrumentation and data processing is described in detail in Ingwersen et al. [2011] and only a summary is presented here. Each station was equipped with a Licor 7500 open path infrared CO₂/H₂O gas analyzer (LI-COR Biosciences Inc., USA) and a CSAT3 3D sonic anemometer (Campbell Scientific Inc., UK). Net radiation was measured with a NR01 4-component sensor (Hukseflux Thermal Sensors, Netherlands) along with air temperature and humidity (HMP45C, Vaisala Inc., USA). Air temperature, humidity, and rainfall were measured on site. Rainfall was measured with a 0.2 mm tipping bucket system at about 1 m height (ARG100, Environmental Measurements LTD, UK). Temperature sensors (Model 107, Campbell Scientific Inc., UK) were installed close to the station, along with TDR probes, and matric potential sensors. Installation depths depended on the solum thickness. At EC1, temperature probes were installed at 2, 6, 15, 30, and 45 cm soil depth and the TDR probes and matric potential sensors at 5, 15, 30, 45, and 75 cm. At EC6, an installation of sensors below 15 cm was not possible because of the shallow solum. Additionally, at both stations three soil heat flux plates (HFP01, Hukseflux Thermal Sensors, Netherlands) were installed 8 cm below ground surface. The EC data were logged at 10 Hz resolution. All other sensor data were stored in 30 min intervals. The EC data were processed using the EC software package TK2 (http://www.bayceer.uni-bayreuth.de/mm/en/software/software/software_dl.php).

Particularly, in heterogeneous landscapes, the energy fluxes measured by EC-technique do not result in a closed energy balance. Therefore, measured turbulent flux data must usually be postprocessed to close the gap. We corrected the EC data assuming that the residual energy entirely consists of sensible heat (H-correction). The theoretical foundation of the H-correction is discussed in Foken [2008]. Further indications that this approach is appropriate for this field site were elaborated and discussed by Ingwersen et al. [2011].

2.2. Simulation Model

2.2.1. Noah-MP

Noah-MP is one of the land-surface components of the Weather Research & Forecasting model (WRF) since version 3.4. It is available as a stand-alone 1-D model (Noah-MP v1.1), which we use in offline mode, i.e., with atmospheric forcing at a temporal resolution of 30 min obtained from field measurements of short-wave and longwave radiation, wind speed, temperature, precipitation, relative humidity, and air pressure. Noah-MP simulates several biophysical and hydrological processes which control LHF and SHF between the canopy and the atmosphere, GHF, and soil water movement. These processes include leaf area development, stomatal conductance, and photosynthesis, surface exchange coefficient for heat, radiation interactions with the vegetation and the soil, as well as the hydrological processes within the canopy and the soil. Noah-MP provides a multiparameterization framework that allows running the model with different combinations of alternative process schemes for individual processes [Niu *et al.*, 2011]. Alternative submodules for 10 physical processes can be applied for up to 4608 different combinations. Soil water fluxes are calculated by the Richards equation using a Campbell/Clapp-Hornberger parameterization of the hydraulic functions [Clapp and Hornberger, 1978]. In the original version of Noah-MP, the root depth and vertical distribution of roots are assumed to be constant during the entire vegetation period.

2.2.2. New Root Growth Option

For testing the impact of considering the dynamics of root growth on simulation results in this study, an additional optional submodule was implemented. This new option, which is a simplification of the root growth module of the crop model SPASS [Wang and Engel, 2000], allows to mimic the increasing extension of the root system during the growth phase of the crop. Root depth and vertical distribution of root density are crucial factors in the sink term of the Richards equation. Thus, the new option directly affects soil moisture simulations. Moreover, root depth is linked to the calculation of the soil moisture factor controlling stomatal resistance (BTR).

The increase of the root depth z_R (cm) is calculated from the maximum root extension rate under optimum conditions, $r_{ext, max}$ (cm d^{-1}), and two reduction factors, $0 \leq f_T \leq 1$ and $0 \leq f_\theta \leq 1$, taking into account the impact of unfavorable temperature and soil moisture conditions in the deepest rooted soil layer:

$$\frac{dz_R}{dt} = r_{ext, max} \cdot f_T(z_R) \cdot f_\theta(z_R) \cdot \left(1 - \frac{z_R}{z_{R, max}}\right) \quad (1)$$

where $z_{R, max}$ (cm) is the maximum rooting depth which is either limited by soil profile depth or ecophysiological constraints of the plant. In this study, $z_{R, max}$ was set to 100 cm and 30 cm at EC1 and EC6, respectively. For $r_{ext, max}$, the default value of 2 cm d^{-1} was used [Wang, 1997]. Moreover, with increasing soil depth an exponential decrease of active root length is assumed in contrast to the uniform distribution of roots over all rooted soil layers used in the original Noah-MP. Details about f_T and f_θ are presented in Appendix A.

2.2.3. Plot-Scale Simulations

Vegetation is represented in Noah-MP by vegetation types which differ in their ecophysiological and hydrological properties. To represent the winter wheat considered in this study, we chose the vegetation type "cropland." Default values of vegetation dynamic and soil parameters are provided in look-up tables. However, using default values can strongly degrade model performance. Consequently, for running the model in the mode that requires predefined values of vegetation dynamics, table values of monthly values of LAI were adjusted to measured time series. Similarly, it has been shown that default values of soil hydraulic properties can yield biased simulations of the dynamics of soil water movement and of plant available water [Gayler *et al.*, 2013]. We therefore replaced the default values of soil hydraulic parameters by values derived from field data by fitting the water retention curve to measured time series of soil moisture and soil matric potential. For running the model in the mode that requires predefined values of vegetation dynamics, table values of monthly values of LAI were adjusted to measured time series. It should be noted, however, that the ensemble simulations were also performed with default soil parameters to analyze potential parameter interdependencies when using different model options. Adjusted and default values of soil parameters as well as LAI dynamics are presented in Table 1. At EC1 (Kraichgau), the depth of the soil column was set to 200 cm in the model, divided in four numerical layers of 10, 30, 60, and 100 cm thickness.

Table 1. Vegetation and Soil Parameters Used in Ensemble Simulations (Default Values in Parenthesis)

Parameter Soil ^a	Description	Units	EC1	EC6
θ_{max}	Max. volumetric soil water	m^3/m^3	0.47 (0.48)	0.48 (0.465)
θ_{wit}	Soil moisture at wilting point	m^3/m^3	0.053 (0.084)	0.123 (0.10)
ψ_{sat}	Saturated soil matric potential	m	0.42 (0.76)	0.23 (0.26)
b	Clapp-Hornberger b		5.0 (5.3)	7.4 (8.2)
Vegetation ^b				
LAI	Leaf area index	m^2/m^2	Apr: 1.61 May: 4.51 Jun: 4.75 Jul: 3.44 Aug: 2.0	Apr: 1.0 May: 2.46 Jun: 4.85 Jul: 4.5 Aug: 2.5

^aDerived from field measurements of soil moisture and matric potential.

^bAdjusted to biweekly field measurements.

At EC6 (Swabian Alb), a soil column of 30 cm depth was simulated, divided in numerical layers of 5, 7, 9, and 9 cm thickness. A differentiation of soil hydraulic properties between the single simulation layers is not possible in Noah-MP.

Simulation runs started on 17 April (EC1) and 1 May (EC6), 1 day after the installation of the meteorological stations. As high-resolution climatic data are required for the forcing of Noah-MP, no simulation of the early growth stages of the wheat crop was possible. Simulations were stopped after harvest of winter wheat (6 August at EC1, 31 August at EC6). To analyze model uncertainty originating from alternative process representations on simulated fluxes in the soil-vegetation-atmosphere continuum, we set up several realizations of Noah-MP by substituting and combining individual process options in different ways. From the processes, for which alternative submodules are available, we selected (1) DEV: dynamic of the aboveground vegetation, (2) CSR: canopy stomatal resistance, (3) BTR: stomata/soil moisture interaction, (4) RUN: surface water infiltration and soil lower boundary for water fluxes, (5) SFC: surface exchange coefficient for heat, (6) RAD: radiation transfer through the vegetation canopy, and (7) RTS: the new option, which considers the dynamic of root growth. Options that offer different ways of addressing frozen soil or snow cover are not considered in this study, because daily mean temperatures did not fall below 0°C during the simulation period. The options used in this study are listed in Table 2. Details on the process descriptions behind the options 1–6 are described in *Niu et al.* [2011]. Altogether, 720 combinations of process parameterizations were investigated. Each simulation run was identically initialized with measured data of soil moisture content and temperature and LAI.

Table 2. Noah-MP Options Investigated in This Study

Process	Options
DEV: Leaf area development	= 1: LAI ^a and GVF ^b predefined (monthly table values) = 2: Prognostic model for LAI, GVF calculated from LAI = 3: LAI predefined as in (1), GVF calculated from LAI
CSR: Canopy stomatal resistance	= 1: Ball-Berry, related to photosynthesis [Ball et al., 1987] = 2: Jarvis-type multiplicative model [Jarvis, 1976]
BTR: Stomata/root/soil moisture interaction	= 1: Function of soil moisture, like Noah [Chen and Dudhia, 2001] = 2: Matric potential related, as in CLM [Oleson et al., 2004] = 3: Matric potential related, as in SSiB [Xue et al., 1991]
RUN: Runoff/soil lower boundary	= 1: TOPMODEL-based runoff/simple groundwater [Niu et al., 2007] = 2: TOPMODEL-based runoff/equilibrium water table [Niu et al., 2005] = 3: Infiltration excess based surface runoff/free drainage [Schaake et al., 1996] = 4: BATS runoff scheme/free drainage [Yang and Dickinson, 1996]
SFC: Surface Exchange Coefficient for Heat	= 1: Based on Monin-Obukhov similarity theory [Brutsaert, 1982] = 2: Neglecting the zero displacement height [Chen et al., 1997]
RAD: Radiation transfer through the vegetation canopy	= 1: Between canopy gap probability depends on 3D canopy structure, maximum gap = 1-GVF when the sun is overhead = 2: Between canopy gap probability = 0 = 3: Between canopy gap probability = 1.0-GVF
RTS: Dynamic Roots	= 1: Static roots (original Noah-MP) = 2: Dynamic roots [Wang, 1997]

^aLAI: Leaf area index.

^bGVF: Green vegetation fraction.

2.2.4. Performance Measures

The performance of the individual model realizations was assessed by comparing simulated fluxes of LHF, SHF, GHF, and soil moisture content in the upper soil layer (SH_2O_{5cm}) with field data. High-resolution eddy flux measurements of LHF and SHF as well as GHF measurements were averaged to weekly mean diurnal cycles of heat fluxes. Standard deviations were calculated to estimate the variability in the measurements. Simulation results were aggregated in the same way. Soil moisture measurements were compared every day at 00:00 h. Nash-Sutcliffe Efficiency (NSE) was used to assess the performance of simulations:

$$NSE = 1 - \frac{\sum (P_i - O_i)^2}{\sum (O_i - \bar{O})^2} \quad (2)$$

where P_i , O_i , and \bar{O} are predictions, observations, and the mean of observations, respectively. NSE can range from $-\infty$ to 1. $NSE = 1$ corresponds to a perfect match between simulated and observed data. $NSE > 0$ indicates that the model predictions are more accurate than modeling the mean of the observed data [Nash and Sutcliffe, 1970]. In addition to the above-mentioned output variables, simulated vertical soil moisture distribution was evaluated by calculating an aggregated NSE of the fits to soil moisture measured at 5, 30, and 75 cm at EC1, and to soil moistures measured at 2.5 cm and 15 cm at EC6.

Substituting one model option for a certain process parameterization with another can have different effects on model accuracy depending on which combination of options are chosen for the remaining processes. To obtain a measure that can be used to compare the impact of a single model option on the overall performance of Noah-MP, for each of the investigated processes $i = DEV, CSR, BTR, RUN, SFC, RAD, \text{ or } RTS$, a weighted mean of the variability in the NSE of LHF, SHF, GHF, and SH_2O , $\overline{\Delta NSE}_i$, was defined (3). This weighted mean of the variability quantifies the shifts in the accuracy of the fits to measured heat fluxes and soil moisture distribution while changing the model options for a certain process, but at the same time running the full ensemble for the remaining processes:

$$\overline{\Delta NSE}_i = \frac{1}{2} \left(\frac{|\overline{\Delta NSE}_{LHF,i}| + |\overline{\Delta NSE}_{SHF,i}| + |\overline{\Delta NSE}_{GHF,i}| + |\overline{\Delta NSE}_{SH_2O,i}|}{3} \right) \quad (3)$$

where $\overline{\Delta NSE}_{LHF,i}$, $\overline{\Delta NSE}_{SHF,i}$, $\overline{\Delta NSE}_{GHF,i}$, and $\overline{\Delta NSE}_{SH_2O,i}$ denote the range within which the mean NSEs of the single observables vary, as the different options for the parameterization of process i are chosen. For example, $\overline{\Delta NSE}_{LHF,CSR}$ is the difference between the mean value of latent heat flux performance of all ensemble members calculating canopy stomatal resistance following "Ball-Berry" and the mean value of latent heat flux performance of the other ensemble members, which calculate canopy stomatal resistance following "Jarvis." In case that three or more options are available for an individual process, the difference between the best and the worst mean NSE is used.

3. Results and Discussion

3.1. Impact of Model Options

In a first step of the analysis, the impact of the submodel choice on simulation results was investigated. The aim of this step was to identify the most critical processes for a successful simultaneous simulation of heat fluxes and soil moisture at both field sites. Moreover, in this step we identified those model options that are obviously not suitable for simulating water and energy fluxes under the given soil and climatic conditions in our field experiments. Inappropriate model options or combinations of model options were subsequently eliminated from the ensemble. The criterion for rejecting an option was a negative mean value of NSE for any of the four observables LHF, SHF, GHF, or SH_2O , averaged over all ensemble members using the respective option.

The $\overline{\Delta NSE}_i$ values for the seven investigated processes are shown for both field sites in Figure 1. Since $\overline{\Delta NSE}_i$ measures only the range within which the mean NSEs of the single observables vary, it cannot be seen from this score whether substituting one option by another results in an increase of the single NSEs or in a decreasing accordance with measurements. For this, the NSEs of the different ensemble runs with the

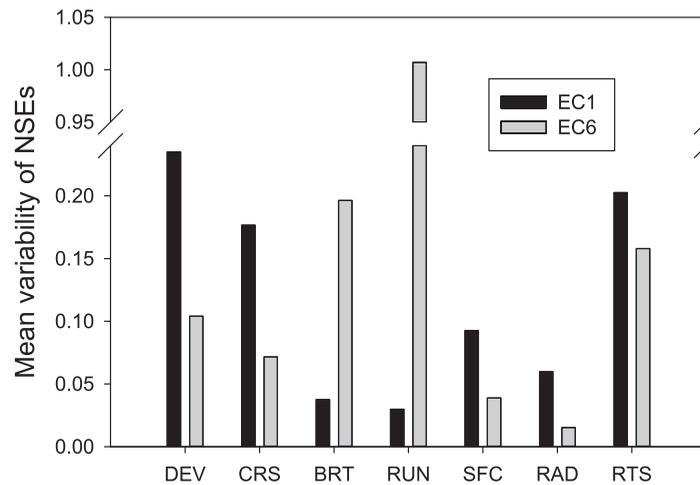


Figure 1. Mean variability in Nash-Sutcliffe Efficiencies achieved by substituting the different Noah-MP options at EC1 (black bars) and EC6 (gray bars), respectively.

respective option activated or not have to be compared. This is exemplarily shown for the two options of RTS in Figure 2, whereas notable effects with respect to the other options are only discussed in section 3.2.

Clear differences in the model behavior between the two field sites can be observed. The most distinctive effects at the deep loess site (EC1) can be observed if options DEV, RTS, and CSR are varied. At EC6, the highest $\overline{\Delta NSE}_i$ values are derived when the different options for RUN, BTR, or RTS are applied. The strong variability of the mean

NSE at EC1 when varying DEV arises mainly from an insufficient simulation of LAI development when the prognostic mode DEV = 2 is activated. Apparently the dynamic vegetation module is not able to accurately capture biomass growth at this managed field. The resulting disagreement between observed and predicted LAI strongly degrades GHF simulations but increases at the same time the fit of simulated soil moisture in the upper 5 cm, which, however, is very low in options DEV = 1 and DEV = 3. Both effects together result in a $\overline{\Delta NSE}_{DEV} = 0.24$. Despite there being only slightly better agreement between observed and predicted LAI under option DEV = 2 at EC6, the impact of poor leaf area simulation on GHF and $SH2O_{5cm}$ is far less pronounced ($\overline{\Delta NSE}_{DEV} = 0.10$). Substituting the multiplicative stomatal resistance model CSR = 2 by Jarvis [1976] with the approach by Ball and Berry [Ball et al., 1987], CSR = 1, increases model accuracy to a greater extent at EC1 ($\overline{\Delta NSE}_{CRS} = 0.18$) compared to EC6 ($\overline{\Delta NSE}_{CRS} = 0.07$). Similarly, a clear positive effect of considering the zero-displacement height in estimating the surface exchange coefficient for heat (SFC = 1) is only given at EC1, mainly due to a strongly improved simulation of GHF.

At EC6, the biggest impact on model accuracy can be observed if option RUN was changed. In particular the “no flux” option RUN = 2 for the lower soil boundary condition results in a very inaccurate simulation of soil moisture dynamics. This is a reasonable model behavior for a shallow soil over karstic bedrock. Best results were obtained with the “free drainage” option RUN = 4. Besides RUN, the option BTR and RTS had the most pronounced effect on simulation results at this field. In the case of BTR, high values of NSE were attained with approaches that simulate the effect of soil water availability on stomatal resistance depending on soil matric potential. Using BTR = 2 or BTR = 3 instead of the soil water content related approach BTR = 1 increases NSE values of $SH2O_{5cm}$ resulting in a $\overline{\Delta NSE}_{BRT}$ of 0.20. Incorporating dynamic root growth has a more pronounced impact on model accuracy at the deeper soil profile in the Kraichgau (EC1) than at the shallow field in the Swabian Alb (EC6) (Figure 2). At EC6, a clear positive effect of the dynamic root model can only be observed for the simulation of the vertical distribution of soil moisture (SH2O). The differences in model behavior are caused by the different depths of the soil profiles. At EC6, when the simulation started (1 May), most of the shallow soil profile had probably already been entered by roots and only a marginal further elongation of the root system was possible during the simulation period, because of the limestone under the soil profile beginning at a depth of 20–30 cm. Consequently, the default Noah-MP assumption of a static root depth was nearly true at this site and only the assumption of an exponential vertical distribution of root length was an advantage of the dynamic root option RTS = 2 compared to the uniform distribution assumed in RTS = 1. In contrast, at EC1 only a small portion of the deep soil profile was entered by roots at the beginning of the simulation period (17 April), and the more realistic simulation of root growth shows a strong positive impact on root water uptake and hence LHF and SHF.

3.2. Analysis of Structural Coherence

Good performance of a model in simulating one or more measured processes may result in poor agreement to data regarding other processes [Wöhling et al., 2013]. Amongst other potential sources of error, such as

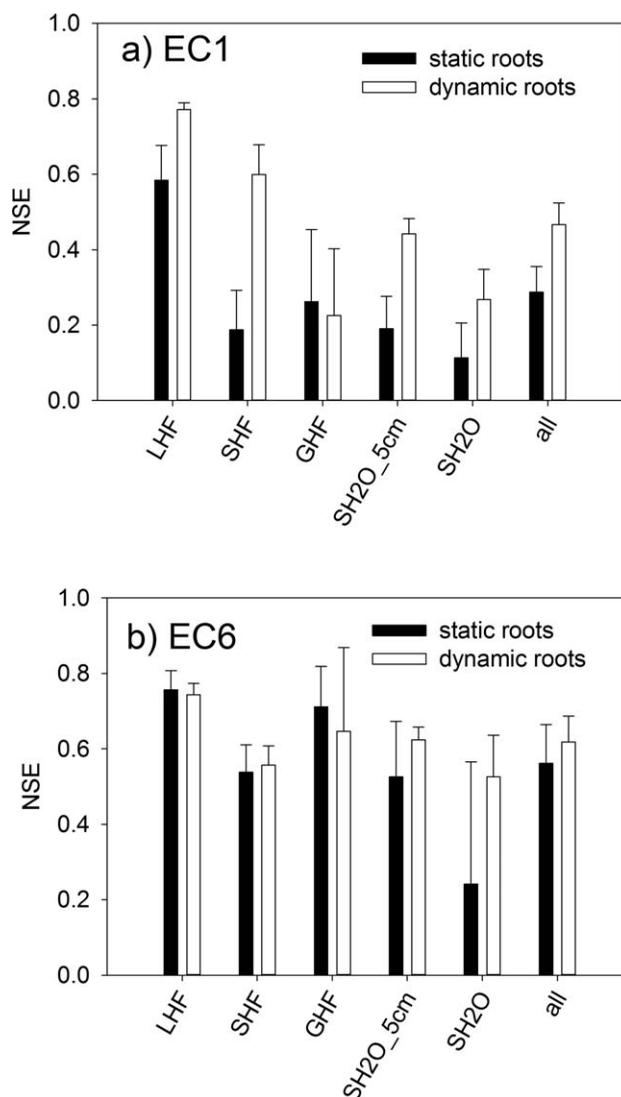


Figure 2. Mean Nash-Sutcliffe efficiencies of model ensembles at (a) EC1 and (b) EC6 without (black bars) and with (white bars) consideration of dynamic root growth for simulated fluxes of latent heat (LHF), sensible heat (SHF), and ground heat (GHF), soil moisture in the upper 5 cm ($SH20_{5cm}$), and soil moisture distribution (SH20).

errors in measurements of parameter values or in boundary conditions, one cause can be structural incoherences in the model or inadequate submodels for single processes. To uncover possible tradeoffs between different state variables in Noah-MP, the ability of the model to simultaneously fit the measurements of LHF, SHF, GHF, $SH20_{5cm}$, and SH20 was evaluated. This step of the model analysis was done with a reduced ensemble, with those options excluded from the full ensemble that obviously result in inadequate simulations of single observables. The excluded options were $DEV = 2$ at EC1 ($NSE_{GHF} = -0.28$), and $RUN = 2$ ($NSE_{SH20} = -1.51$) and $RUN = 3$ ($NSE_{SH20} = -0.34$) at EC6, resulting in ensemble sizes of 576 at EC1 and 360 at EC6. Cross comparisons of the observable-specific NSEs for the individual ensemble members are summarized for EC1 in Figure 3 and for EC6 in Figure 4. The single subplots of these figures show the relationships between the NSEs of the individual ensemble members of each two output variables. As additional information, the symbols in the plots are differentiated by color corresponding to the fit of the respective simulation run to the measured vertical soil moisture distribution. Green, blue, and red symbols indicate the simulations with highest, medium, and lowest values of NSE_{SH20} , respectively.

Similar, but quite differently pronounced, patterns of the relationships between the individual NSE values can be observed at both contrasting field sites. Combinations of model options that achieve a good result for LHF also have high values of NSE_{SHF} . The range within which the values of NSE_{LHF} and NSE_{SHF} vary, however, is much greater at EC1 than at EC6. At least at EC1, there is also a certain relationship between NSE_{SH20} and successful simulations of latent and sensible heat. Some of the ensemble members, however, achieve high values of NSE_{LHF} and NSE_{SHF} concurrent with poor agreement between measured and observed vertical soil moisture distribution. The new option $RTS = 2$ simultaneously has a significant positive effect on LHF, SHF, and SH20, but not on GHF. At EC6, there is nearly no impact of soil moisture simulation on the results achieved for LHF and SHF. Obviously, during the simulation period, the energy fluxes between land surface and atmosphere were more governed by radiation than by soil moisture, which can be explained by the high water holding capacity of the loess soil at EC1 and the sufficient precipitation during the vegetation period at EC6 (494 mm from 1 May to 31 August). The NSE values achieved for both $SH20_{5cm}$ and SH20 are significantly higher at EC6 than at EC1. The reason for this lies in the constraint of Noah-MP that

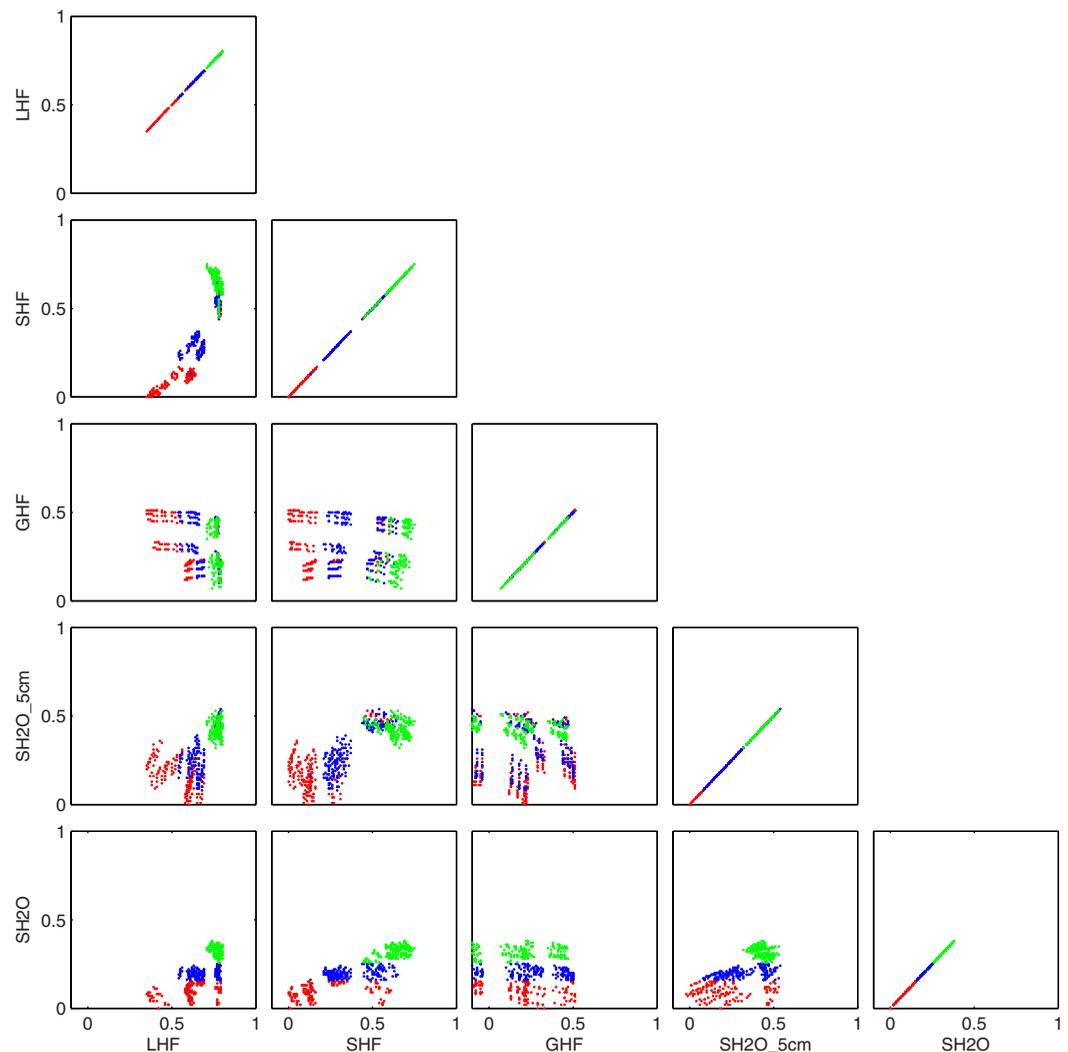


Figure 3. Correlation between Nash-Sutcliffe efficiencies of latent heat flux (LHF), sensible heat flux (SHF), ground heat flux (GHF), soil moisture in the upper 5 cm (SH2O_5cm), and soil moisture distribution (SH2O) at the Kraichgau field site EC1. Different colors of symbols indicate ensemble members showing best (green), medium (blue), and lowest fits to measured soil moisture distribution.

the soil profile is assumed to be homogeneous with regard to its hydraulic functions. Autonomous sets of parameters for the individual soil layers are not possible in this model. Consequently, at EC1, where the soil profile is much deeper compared to EC6, there is greater necessity to compromise between the differing individual relationships of soil moisture and matric potential measured at different soil depth (see section 3.5).

At both field sites, the adequacy of GHF simulations is nearly independent from the accuracy with which the other output variables are simulated. Amongst the highest values of NSE_{GHF} are ensemble members which show good, medium, or poor agreement with measurements of soil moisture, LHF, or SHF. Specific combinations of model options for successful simulation of GHF at EC1 require $SFC = 1$ together with $RAD = 2$. Including $SFC = 2$, $RAD = 1$, or $RAD = 3$ are responsible for a splitting of the scatterplots into different branches. The same can be observed at EC6, however, this effect is much less pronounced. Moreover, at EC6, some of the simulations with high values of NSE_{SH2O} produce very weak simulations of GHF. This is the case if the prognostic leaf area module $DEV = 2$ is chosen together with the new root module $RTS = 2$. Obviously, in the case of GHF predictions, the more simplified option $RTS = 1$ compensates the insufficient LAI simulations. In contrast, $RTS = 2$ has a slightly positive effect on LHF, SHF, and SH2O simulation results. At both field sites, the strongest compromise between individual model realizations must be made when

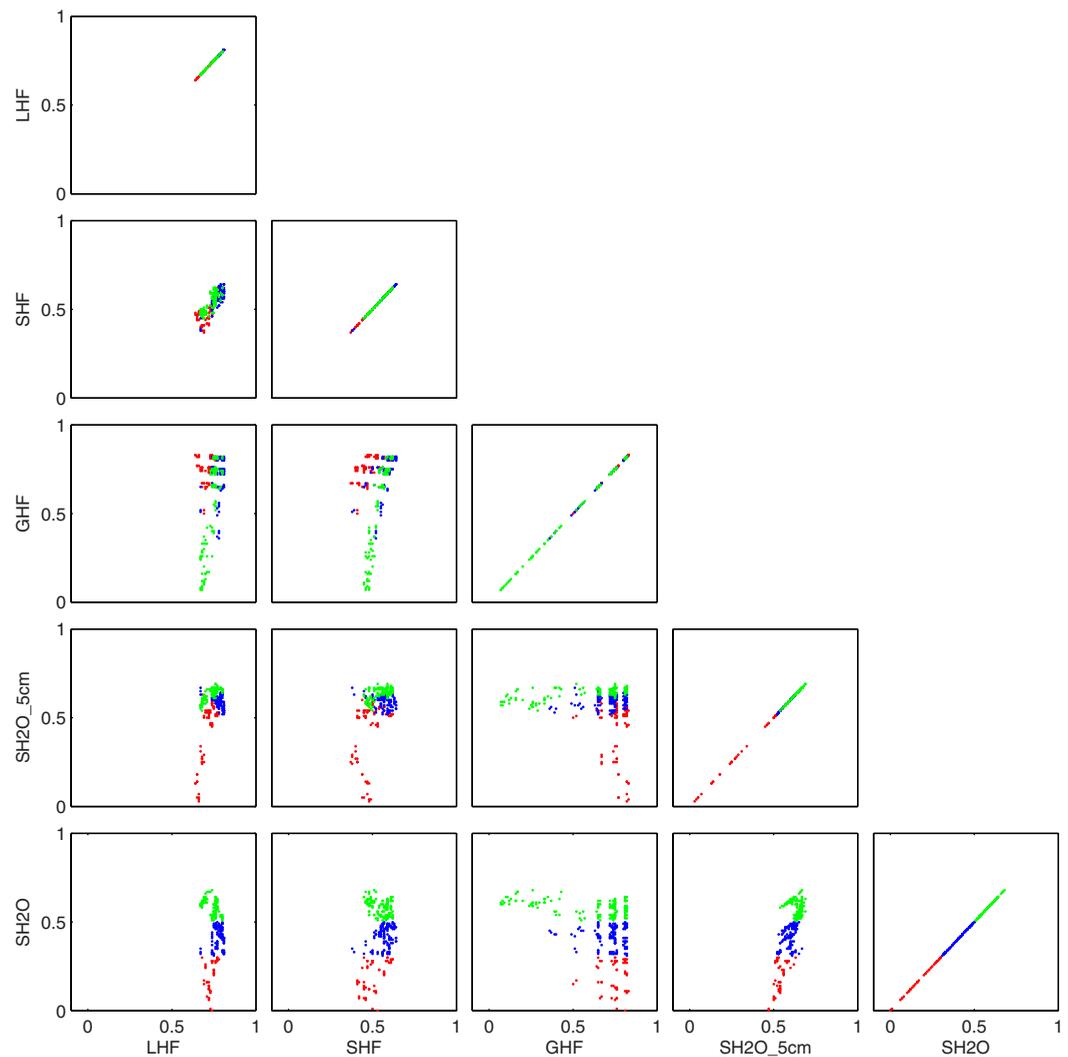


Figure 4. Correlation between Nash-Sutcliffe efficiencies of latent heat flux (LHF), sensible heat flux (SHF), ground heat flux (GHF), soil moisture in the upper 5 cm (SH2O_5cm), and soil moisture distribution (SH2O) at the Swabian Alb field site EC6. Different colors of symbols indicate ensemble members showing best (green), medium (blue), and lowest fits to measured soil moisture distribution.

optimizing GHF simulations, because the highest values of NSE_{GHF} go along with weak NSE values of the other output variables. This behavior of Noah-MP is a clear indication of structural incoherences in the model in relation to GHF simulations.

3.3. Transferability Between Field Sites

Combinations of model options that result in high NSE values at one field do not necessarily result in accurate simulations at another, particularly if soil and climatic conditions are different. Likewise, substituting a submodule that improves simulation accuracy at one site may decrease the same at another. We therefore investigated whether or not ensemble members that provide good results at one field site also do so at the other location of our study. We therefore calculated averaged NSEs, $NSE_{all} = \frac{1}{4}(NSE_{LHF} + NSE_{SHF} + NSE_{GHF} + NSE_{SH2O})$, for the individual ensemble members and compared their values obtained at EC1 and EC6. In this exercise, the ensemble size was reduced to 288 considering only those combinations of model options that were previously not excluded at either of the two field sites. Figure 5 shows the relationship between the averaged model performances at EC1 and EC6 for different combinations of model options. Combinations of model options, which are most successful at the one field site, are also among the most successful model realizations at the other field site. Considering the dynamics of root growth increases the performance of the model ensemble to match measured energy and water

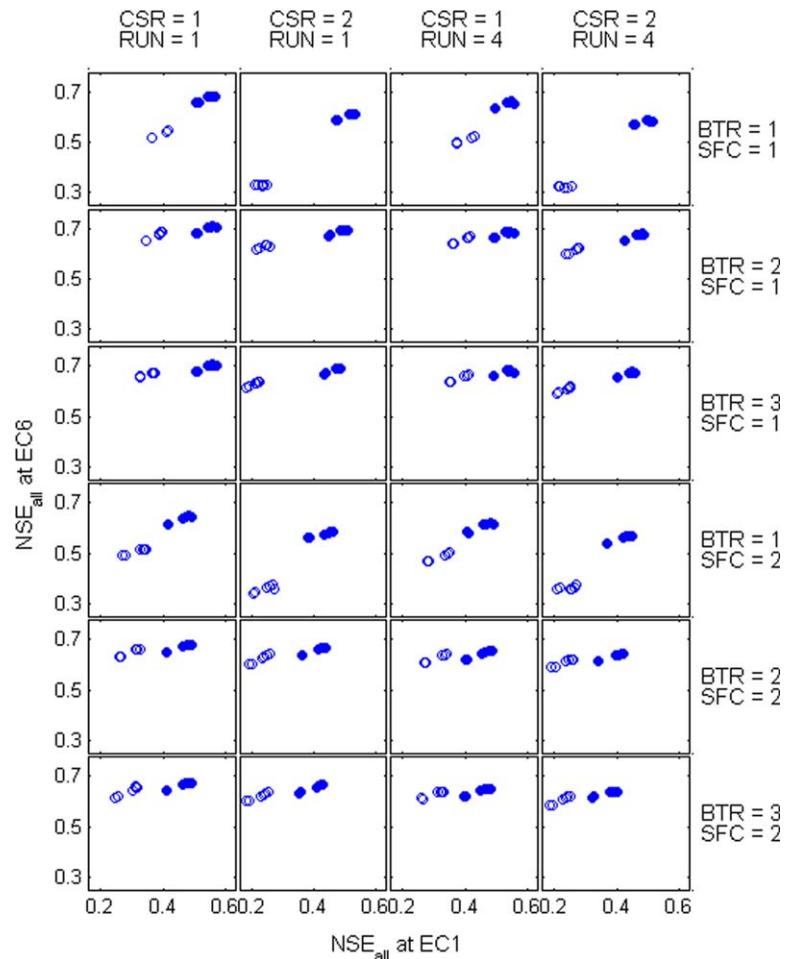


Figure 5. Relationship between averaged model performances (NSE_{all}) at EC1 and EC6 for different combinations of model options. Each subplot represents a 12 member ensemble with $DEV = 1$ or 3 , $RAD = 1, 2$, or 3 , $RTS = 1$ or 2 , and the other model options defined in column and row headers. Open circles: Noah-MP realizations with $RTS = 1$; closed circles: Noah-MP realizations with $RTS = 2$. NSE_{all} integrates NSEs of latent heat flux (LHF), sensible heat flux (SHF), ground heat flux (GHF), and soil moisture distribution (SH2O).

fluxes at both field sites (indicated by the closed symbols in Figure 5). However, the benefit of this option strongly depends on the model options used for the other processes and on the field site. $RTS = 2$ increases model performance at EC1 for all ensemble members. At EC6, a significant increase of NSE_{all} when using $RTS = 2$ instead of $RTS = 1$ is only given for $BTR = 1$. Interactions and field site dependencies can also be observed between other physics options. Changing from $SFC = 2$ to $SFC = 1$ increases the model performance at EC1 mainly if $RTS = 2$ is selected. $BTR = 1$ decreases NSE_{all} at EC6 (mainly if $CSR = 2$) but not at EC1. However, the adverse effect of $BTR = 1$ diminishes if $RTS = 2$. Choosing $CSR = 1$ instead of $CSR = 2$ increases model performance at both sites, if $BTR = 1$, but only at EC1, if $BTR = 2$ or 3 .

A number of 25 simulations simultaneously achieve $NSE_{all} > 0.50$ at EC1 and $NSE_{all} > 0.65$ at EC6. From these most successful simulations, 20 belong to an ensemble with $DEV = 1$ or 3 , $CSR = 1$, $BTR = 2$ or 3 , $RUN = 1$ or 4 , $SFC = 1$, $RAD = 1, 2$ or 3 , $RTS = 2$, which we therefore consider as the most appropriate in our study (see Table 3). Thus, taking into account some site specific characteristics such as the lower soil boundary conditions, the transferability of the findings from one location to the other was shown in this experiment.

3.4. Interaction of Parameter Values and Multiphysics Options

The focus of the study was on model uncertainty that arise from alternative representations of individual model components. However, this is only one of different sources of uncertainty in land-surface models. Errors in measurements of input data and system responses, uncertainty in the values of parameters used

Table 3. Mean Values \pm Standard Efficiencies of Nash-Sutcliffe Efficiencies of the Ensemble (24 Members) Showing Best Performance at EC1 and EC6 (DEV = 1, 3; CSR = 1; BTR = 2, 3; RUN = 1, 4; SFC = 1; RAD = 1, 2, 3; RTS = 2)

	EC1	EC6
NSE _{LHF}	0.76 \pm 0.01	0.76 \pm 0.01
NSE _{SHF}	0.70 \pm 0.01	0.61 \pm 0.01
NSE _{GHF}	0.35 \pm 0.09	0.79 \pm 0.03
NSE _{SH2O_5 cm}	0.44 \pm 0.04	0.65 \pm 0.02
NSE _{SH2O}	0.32 \pm 0.03	0.59 \pm 0.04
NSE _{all}	0.53 \pm 0.03	0.69 \pm 0.02

in the individual representations of physical processes, ambiguities in the overall model structure, and the interplay of all these components also contribute to the total uncertainty of model predictions [e.g., Clark et al., 2011a]. Taking all these possible error sources into account when searching for the

most appropriate model structure for a given study objective is quite a challenging task. In particular, inadequate parameter values in individual submodules, time-dependent parameters or incomplete understanding of physical processes can hamper the evaluation of different schemes. Although some of these issues are still not completely resolved and there is an ongoing discussion about the most appropriate approaches to tackle them [Clark et al., 2011a, 2012; Beven et al., 2012], various formal frameworks have been developed and applied in recent years to test individual model hypotheses (the representation of physical processes) and to reject models or submodules in complex system models [Beven and Freer, 2001; Clark et al., 2008; Gupta et al., 2008; Clark et al., 2011b; Fenicia et al., 2011; Gupta et al., 2012; Fenicia et al., 2014]. Though predominantly applied in catchment hydrology, these frameworks should also be transferable to other fields of environmental modeling [Clark et al., 2011a]. However, implementing a formal testing framework is very difficult and time consuming in the context of the complex land-surface model Noah-MP and could not be attempted within the scope of this study. Nevertheless, also in the simple but structured model selection strategy of the presented study, the aspect of interactions between parameter values and alternative model structures could be considered as subsequently explained.

The interplay of parameter values and alternative model structures can play an important role in model the performance of model simulations and consequently also in the submodule selection process. Previous research has shown that optimal parameter values for individual submodels vary depending on the choice of submodels for other processes or on the combination of process representation. Wöhling et al. [2013] recently investigated the impact of structural model complexity on the performance of five coupled soil-plant models and the Community Land Model (CLM 3.5) to simultaneously simulate soil moisture, evapotranspiration, and leaf area index. In this study, five different crop modules, covering a wide range in the details of representing crop growth, were coupled to identical soil modules. A large variability in the corresponding soil hydraulic parameter sets that optimize the performance of the individual model realizations was found. Likewise, in a study with three versions of an earlier version of the Noah model, Rosero et al. [2010] have found differences in the sensitivity of parameters of submodules and in optimal submodule parameters between several combinations of model schemes.

Consequently, the optimal combination of physics options as reported in the previous sections may also depend on the choice of parameter values used in individual submodules. We therefore repeated the analysis of the transferability of findings between EC1 and EC6 (section 3.3) using the default values of soil hydraulic parameters provided in the lookup tables of Noah-MP instead of the parameter values derived from field measurements (Table 1). These parameters directly affect soil water movement but also the stomata-root-soil moisture control BTR [Niu et al., 2011]. This analysis was conducted to exemplarily investigate the interplay between parameter values and model structure in carving out the best combinations of process representations in our study. The results are summarized in Figure 6 and are subsequently discussed. Using default parameters strongly degraded model performance at both field sites. At EC1, the average of NSE_{all} decreases from 0.38 to 0.34 at EC6, and from 0.61 to 0.46 at EC6. This is mainly due to a significant degradation in the performance of soil moisture simulations, whereas heat flux simulations were much less affected. The degrading effect of using the default soil parameters is much less pronounced, if RTS = 2 is selected. In particular, in this case the split pattern (originating from the strong negative effect of choosing BTR = 1 on simulation results at EC6), is not visible if default soil parameters are used and RTS = 2. This example demonstrates how the discovery of structural deficits in a model can depend on both, a suitable combination of physical process representations and the parameter values used in individual submodules.

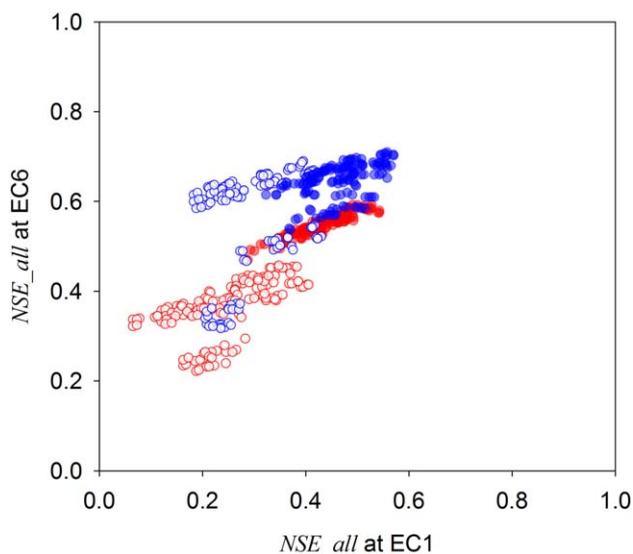


Figure 6. Summarized relationship between averaged model performance (NSE_{all}) at EC1 and EC6 using default (red) and adjusted (blue) values for soil hydraulic parameters. Open circles: Noah-MP realizations with $RTS = 1$; closed circles: Noah-MP realizations with $RTS = 2$.

In summary of the above, the main findings of our study are only marginally affected by the two different sets of soil parameters. We are aware, though, that this example can only provide limited insights into the complex interplay of parameter values and submodel choices. A more comprehensive hypothesis testing framework including optimization of parameters for each of the submodule combinations goes beyond the scope of this study. Moreover, for practical applications involving ensembles with a large number of multiphysics model realizations, the required data to meaningfully constrain the parameters and to analyze parameter correlations in the various submodules are typically not available at the required level of precision and the scale of interest, which is particularly the case in large-scale coupled land-surface schemes.

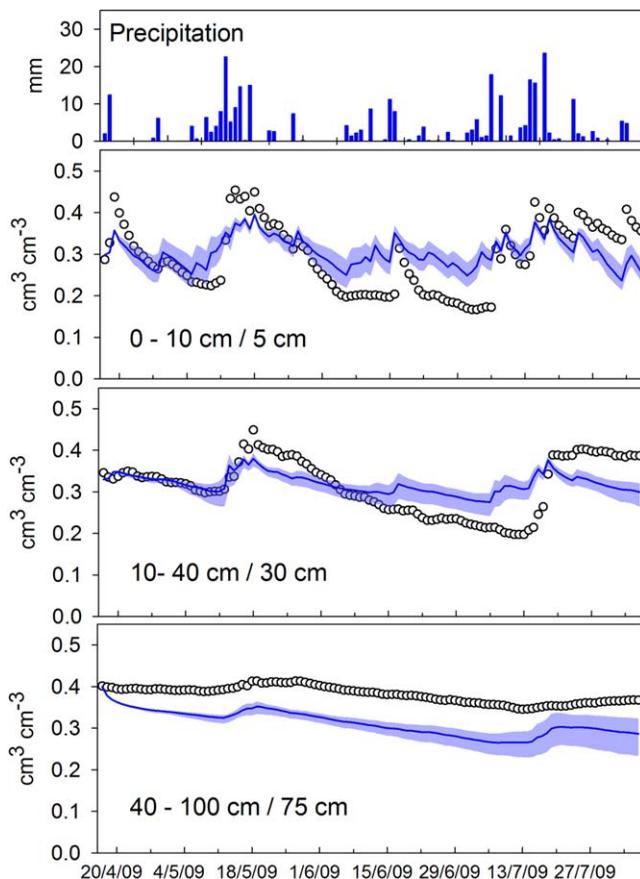


Figure 7. Ensemble means (solid lines) and prediction uncertainties (5–95% quantile range, shaded regions) of simulated soil moisture dynamics at EC1 at 0–10 cm, 10–40 cm, and 40–100 cm depth, simulated by the reduced ensemble (288 members). Observations, symbolized by circles, were made in 5, 30, and 75 cm, respectively.

3.5. Ensemble Simulations

Finally, the uncertainty in model predictions, which originates from the alternative modeling approaches used for the different submodules was estimated by the mean values and the 5% and 95% quantiles of soil moisture and heat flux ensemble simulations. Since a range of different, nonoptimal but plausible parameterization schemes can nevertheless provide meaningful results, in particular if the “optimal” configuration may vary with time (e.g., between early development stages and the period of senescence), ensemble simulations with the 288 members described in the earlier sections were conducted. For the soil parameters, the adjusted values derived from field measurements were taken. The results of the ensemble simulations are shown in Figures 7–10. At EC1, the high

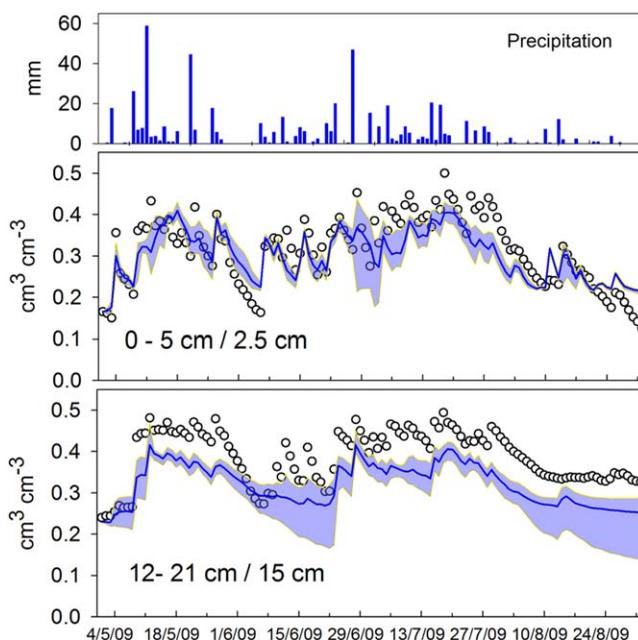


Figure 8. Ensemble means (solid lines) and prediction uncertainties (5–95% quantile range, shaded regions) of soil moisture dynamics at EC6 at 0–5 cm and 12–21 cm depth, simulated by the reduced ensemble (288 members). Observations, symbolized by circles, were made in 2.5 and 15 cm, respectively.

variability in soil moisture observations in the upper soil layer is clearly underestimated by the model, whereas the dynamic in the deepest soil layer is overestimated. This is possibly caused by the “mean parameters” that have to be used for the whole soil column without differentiation between individual soil horizons (see section 3.2. and Figure 7). The uncertainty in soil moisture simulations at 5 cm depth is most pronounced during dry periods, whereas prolonged precipitation events narrow the individual simulation runs to near saturation water contents. While in the upper soil layer the increase in model uncertainty is probably caused by differences in evapotranspiration processes, which take place at small time scales, the increasing uncertainty at 75 cm depth results from an

accumulation of small differences in simulated soil water fluxes. At EC6, soil moisture dynamics in both layers are better simulated than at EC1 (Figure 8). Obviously, the more shallow soil profile at this field site can be better represented by a single set of hydraulic parameters. However, absolute values of soil moisture in the lower soil layer are mostly underestimated, indicating a simulation bias toward excessive drainage. In this layer, a similar pattern of model uncertainty as in the upper layer of EC1 can be observed. The long drying out of the soil at the end of the vegetation period reduces the uncertainty in the upper 5 cm but leads to an increased uncertainty in soil moisture in 12–21 cm depth.

During most of the time, the seasonal dynamic of weekly averaged diurnal cycles of LHF and, with some restrictions, SHF, can be well simulated at both field sites (Figures 9 and 10). Exceptions are the last 3 weeks of the growing season at EC1, where plants became senescent. During this period, LHF, which is associated with water uptake by vegetation, is clearly overestimated while SHF is underestimated. At EC6, in contrast, LHF and SHF are better matched during senescence. The divergent model behavior at both locations results from differences in the amount of plant available soil water at the end of the vegetation period. At the Kraichgau site with deep loess soils (EC1), much more water is stored in the soil compared to the Swabian Alb location (EC6), where soil depth is only 20–30 cm over a karstic underground and the soil moisture decreased rapidly at the end of the vegetation period. Reduced evapotranspiration rates during dry soil periods are well simulated by Noah-MP, but a reduction of transpiration rates, which is caused by a strong degeneration of the plant hydraulic system during senescence, cannot be reproduced by the model. This corroborates results obtained by Gayler *et al.* [2013] in a study on the performance of the community land model CLM 3.5 compared to simulation results obtained from a crop model with an explicit representation of root senescence, and by Wöhling *et al.* [2013] in a multimodel study with four additional soil-plant models of different complexity. A similar shortcoming concerning vegetation dynamics was also discussed in a study by Ingwersen *et al.* [2011] with the land-surface model Noah [Chen and Dudhia, 2001]. Uncertainty bands around the ensemble mean of LHF and SHF are more pronounced at EC1 than at EC6. This results mainly from the stronger effect of modeled root growth dynamics on turbulent fluxes at the deep soil location (compare 2).

At EC1, GHF is clearly overestimated in the first half of the day during most of the vegetation period (Figure 9). The same can be observed during the last third of the vegetation period at EC6 (Figure 10). The extent of the deviations between simulations and measurements depends only slightly on the chosen combination

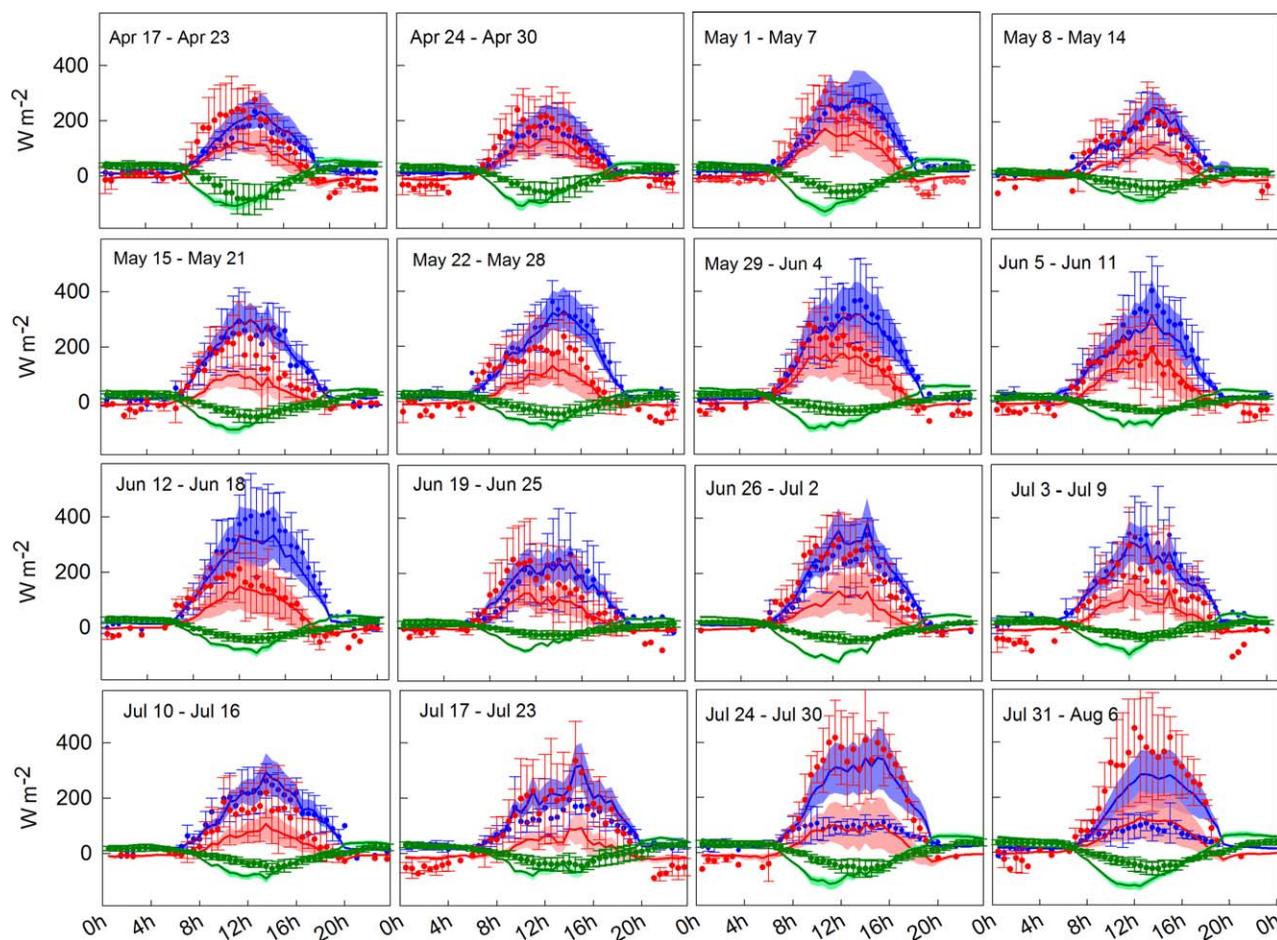


Figure 9. Weekly averaged ensemble simulations (288 members) of diurnal cycles of latent heat (LHF, blue), sensible heat (SHF, red), and ground heat (GHF, green) fluxes during the growing season 2009 (17 April–6 August at EC1 (Kraichgau) together with observed values (symbols + error bars, same colors). Solid lines represent ensemble means, shaded regions denote the 5–95% quantile range of the ensemble simulations. Error bars symbolize standard deviations resulting from the weekly averaging of observations.

of submodules, which can be seen on the narrow uncertainty bands around the ensemble mean simulations. Moreover, the simulation results are hardly sensitive to the variation of the soil thermal and hydraulic parameters used in the soil heat transport equations (not presented here). Thus, the poor performance of Noah-MP with respect to the GHF simulation in this study cannot be attributed to inadequate parameter values. Rather, it shows again the structural deficits of the model with respect to GHF.

4. Summary and Conclusions

Many alternative equations are proposed to simulate the biogeophysical processes in the soil-vegetation-atmosphere, which are most relevant for the exchange of water and energy between the land surface and the atmosphere. In general, it cannot be decided in advance, which process parameterizations or which combination of process parameterizations are most appropriate for a given simulation purpose. Consequently, in estimating the uncertainty of simulation results of LSMs, the uncertainty inherent in alternative parameterization options must be considered as well as the uncertainty in parameter values. In this study, we used the community land-surface model Noah-MP, which offers a multioption framework with a multitude of exchangeable process representations, to find the most appropriate process parameterizations for simultaneously simulating soil moisture, ground heat flux as well as the exchange of latent and sensible heat between the land surface and atmosphere at two contrasting winter wheat fields in South-West Germany. To acknowledge the dynamics of root growth in this study, we enhanced the multioptions

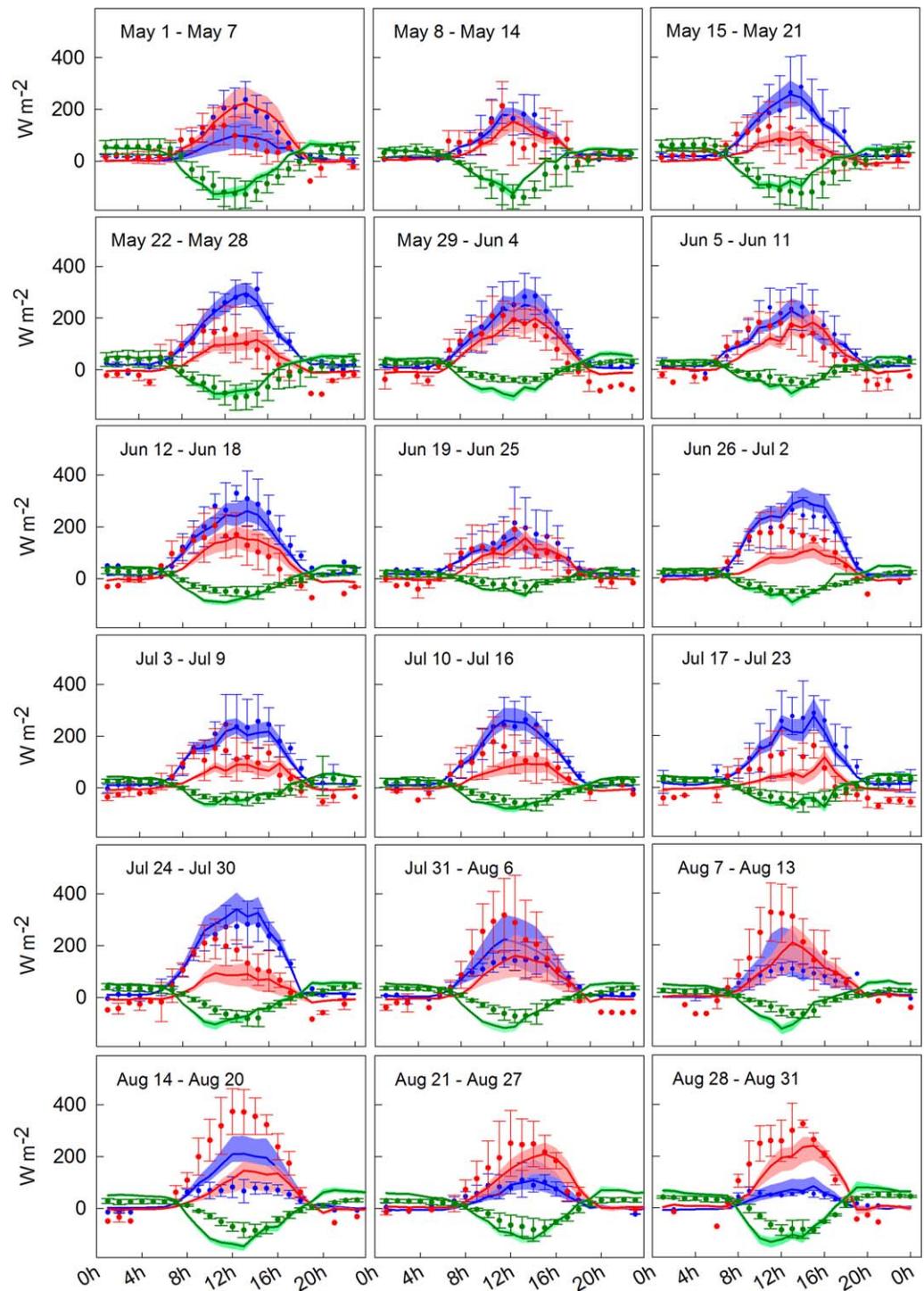


Figure 10. Weekly averaged ensemble simulations (288 members) of diurnal cycles of latent heat (LHF, blue), sensible heat (SHF, red), and ground heat (GHF, green) fluxes during the growing season 2009 (1 May–31 August) at EC6 (Swabian Alb) together with observed values (symbols + error bars, same colors). Solid lines represent ensemble means, shaded regions denote the 5–95% quantile range of the ensemble simulations. Error bars symbolize standard deviations resulting from the weekly averaging of observations.

framework of Noah-MP with a new submodule adapted from the agricultural crop model SPASS [Wang, 1997]. Model uncertainty originating from alternatives in representing physical processes was analyzed via a multiphysics ensemble approach. Moreover, inconsistencies in the model structure and the transferability of the findings from one location to the other were investigated.

Our results show that substituting individual process representations can strongly affect simulation results. The newly implemented option that allows for the simulation of root growth of crops significantly improved the performance of the model. However, the way Noah-MP responds to the alternative model options is quite different between both field sites investigated in this study. This can be partly attributed to differences in soil properties and soil profile depth at each site. In particular, an inappropriate representation of the lower boundary condition in the soil water flux module generated faulty results at the Swabian Alb site where the soil profile is very shallow. This has strong implications for large-scale applications of LSMs, because information about soil profile depth is much too imprecise or even missing in soil maps typically used in such studies. Likewise, the differences in model response to considering dynamic of roots originates from differences in soil depth.

There is no single combination of model options or “best” model, which optimizes all of the considered output variables at the same time. An increasing accuracy of one or more individual output variables by choosing the most appropriate process parameterization with respect to these variables can be accompanied by decreased accuracy in other output variables. Consequently, a compromise between the individual observables must be found. The extent of the compromises between latent heat, sensible heat, ground heat, and soil moisture in our study can be seen from the shape of the scatterplots presented in Figures 3 and 4. For example, combinations of model options can be found which simultaneously increase the accuracy of latent and sensible heat flux. Such correlations cannot be found for ground heat flux. Moreover, improving ground heat flux comes at the expense of soil moisture simulation accuracy, indicating structural deficits in the model.

Inaccuracies in soil moisture simulations are probably caused by overly simplified representation of the soil hydraulic properties in Noah-MP, which allows no differentiation between individual soil horizons. This constraint has more adverse implications regarding the deep soil profile at the Kraichgau site than the shallow soil at Swabian Alb. Augmenting the present routine for soil water movement with a more detailed approach that allows depth-dependent parameterization of the hydraulic functions would probably be one of the most promising approaches to enhancing the performance of Noah-MP, at least in small-scale studies, where the required data are available. Nevertheless, in certain situations Noah-MP can provide good predictions of heat fluxes without necessarily predicting soil moisture accurately. Apparently, there are combinations of model options that compensate structural deficits in soil water flux simulations.

Excluding those process parameterizations that are not appropriate for the given site conditions results in a multiphysics ensemble that is highly transferable between both field sites. In particular, if the newly introduced option “dynamic roots” is activated, ensemble members achieve good performance in simultaneously simulating heat fluxes and soil moisture at both field sites. We therefore conclude that considering root growth dynamics is an important factor at agricultural sites, where annual plants are cultivated. The robustness of our conclusions was demonstrated, when using default values of soil parameters instead of values derived from field measurements. Although the model performance degraded at both field sites, the main findings of our study are only marginally affected.

At least at the Swabian Alb, clear improvements of simulating vertical soil moisture distribution could also be achieved by substituting the water content-based approach with a matric potential-based approach for estimating the soil moisture/stomata interaction. This process is closely related to root water uptake and the latter type of parameterization can be seen as more realistic in the sense of root physiology, suggesting the presumption that more biological realism may improve the performance of LSMs. Insufficient simulation results were achieved at both locations using the prognostic mode for leaf area development, which was therefore excluded from the multiphysics ensemble. Clearly, the dynamic submodel of Noah-MP is not able to adequately simulate phenology of winter cereals, and hence biomass growth, in managed agro-ecosystems in regions where such crops are prevalent. However, reliable simulation of vegetation responses to changing temperatures and precipitation is essential if LSMs are used for projections in future climate conditions.

Thus, aside from a more detailed representation of soil water movement, a more “biological” representation of vegetation-related processes seems to be the most important control for consistent simulation of energy and water fluxes in the soil-plant system. In particular, further advances in the representation of leaf area dynamics and of root/soil moisture interactions seem to be the most promising starting points for improving the simulation of feedbacks between underground, land surface, and atmosphere in fully coupled

hydrological and atmospheric models. It should, however, be considered that changes in the land-surface model may have considerable impacts on the coupling with the atmosphere via parameterizations of the atmospheric surface layer. A full assessment of the impact is only possible by testing the performance of fully coupled simulations which will be the next natural step after demonstrating improvements of the land-surface model.

Appendix A: Reduction Functions of Root Depth Growth

Wang [1997] gives the following soil temperature (f_T) and soil moisture (f_θ) dependencies of root elongation:

$$f_T(z) = \frac{2(T(z) - T_{min})^\alpha (T_{opt} - T_{min})^\alpha - (T(z) - T_{min})^{2\alpha}}{(T_{opt} - T_{min})^{2\alpha}}$$

$$\text{where } \alpha = \ln 2 \ln \left[(T_{max} - T_{min}) / (T_{opt} - T_{min}) \right]$$

and

$$f_\theta(z) = \min \left\{ 1, 4 \cdot \frac{\theta(z) - \theta_{pwp}}{\theta_{fc} - \theta_{pwp}} \right\}$$

Here z is the depth of the root system. For wheat the three cardinal temperatures T_{min} , T_{opt} , and T_{max} are 0, 25, and 35°C, respectively. When the relative extractable soil water content in the deepest root layer is less than 25% of the total extractable soil water, the root extension rate will be reduced linearly.

Acknowledgments

This work was supported by a grant from the Ministry of Science, Research and Arts of Baden-Württemberg (AZ Zu 33-721.3-2), the Helmholtz Center for Environmental Research, Leipzig (UFZ), and by the German Research Foundation (DFG) in the frame of the integrated research project FOR 1695 "Agricultural Landscapes under Global Climate Change—Processes and Feedbacks on a Regional Scale."

References

- Ajami, N. K., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res.*, *43*, W01403, doi:10.1029/2005WR004745.
- Asseng, S., et al. (2013), Uncertainty in simulating wheat yields under climate change, *Nat. Clim. Change*, *3*, 827–832.
- Aurbacher, J., P. S. Parker, G. A. Calberto Sánchez, J. Steinbach, E. Reinmuth, J. Ingwersen, and S. Dabbert (2013), Influence of climate change on short term management of field crops—A modelling approach, *Agric. Syst.*, *119*, 44–57.
- Ball, M., I. Woodrow, and J. Berry (1987), A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in *Progress in Photosynthesis Research*, edited by J. Biggins, pp. 221–224, Martinus Nijhoff, Dordrecht, Netherlands.
- Beven, K., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, *249*, 11–29.
- Beven, K., P. Smith, I. Westerberg, and J. Freer (2012), Comment on "Pursuing the method of multiple working hypotheses for hydrological modeling" by P. Clark et al., *Water Resour. Res.*, *48*, W11801, doi:10.1029/2012WR012282.
- Bonan, G. B., P. J. Lawrence, K. W. Oleson, S. Levis, M. Jung, M. Reichstein, D. M. Lawrence, and S. C. Swenson (2011), Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, *J. Geophys. Res.*, *116*, G02014, doi:10.1029/2010JG001593.
- Breuer, H., F. Ács, B. Laza, A. Horváth, I. Matyasovszky, and K. Rajkai (2012), Sensitivity of MMS-simulated planetary boundary layer height to soil dataset: Comparison of soil and atmospheric effects, *Theor. Appl. Climatol.*, *109*(3–4), 577–590.
- Brutsaert, W. A. (1982), *Evaporation into the Atmosphere*, D. Reidel, Dordrecht, Netherlands.
- Challinor, A. J., M. S. Smith, and P. Thornton (2013), Use of agro-climate ensembles for quantifying uncertainty and informing adaptation, *Agric. For. Meteorol.*, *170*, 2–7.
- Chen, F., and J. Dudhia (2001), Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, *Mon. Weather Rev.*, *129*(4), 569–585.
- Chen, F., Z. Janjić, and K. Mitchell (1997), Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP Mesoscale Eta model, *Boundary Layer Meteorol.*, *85*(3), 391–421.
- Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, *14*(4), 601–604.
- Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.
- Clark, M. P., D. Kavetski, and F. Fenicia (2011a), Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resour. Res.*, *47*, W09301, doi:10.1029/2010WR009827.
- Clark, M. P., H. K. McMillan, D. B. G. Collins, D. Kavetski, and R. A. Woods (2011b), Hydrological field data from a modeller's perspective: Part 2: Process-based evaluation of model hypotheses, *Hydrol. Processes*, *25*(4), 523–543.
- Clark, M. P., D. Kavetski, and F. Fenicia (2012), Reply to comment by K. Beven et al. on "Pursuing the method of multiple working hypotheses for hydrological modeling," *Water Resour. Res.*, *48*, W11802, doi:10.1029/2012WR012547.
- Dirmeyer, P. A. (2013), Characteristics of the water cycle and land-atmosphere interactions from a comprehensive reforecast and reanalysis data set: CFSv2, *Clim. Dyn.*, *41*, 1083–1097.

- Dirmeyer, P. A., R. D. Koster, and Z. Guo (2006), Do global models properly represent the feedback between land and atmosphere?, *J. Hydrometeorol.*, *7*, 1177–1198.
- Fenicia, F., D. Kavetski, and H. H. G. Savenije (2011), Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, *Water Resour. Res.*, *47*, W11510, doi:10.1029/2010WR010174.
- Fenicia, F., D. Kavetski, H. H. G. Savenije, M. P. Clark, G. Schoups, L. Pfister, and J. Freer (2014), Catchment properties, function, and conceptual model representation: Is there a correspondence?, *Hydrol. Processes*, *28*(4), 2451–2467.
- Foken, T. (2008), The energy balance closure problem: An overview, *Ecol. Appl.*, *18*, 1351–1367.
- Gayler, S., J. Ingwersen, E. Priesack, T. Wöhling, V. Wulfmeyer, and T. Streck (2013), Assessing the relevance of sub surface processes for the simulation of evapotranspiration and soil moisture dynamics with CLM3.5: Comparison with field data and crop model simulations, *Environ. Earth Sci.*, *69*(2), 415–427.
- Grasselt, R., D. Schüttemeyer, K. Warrach-Sagi, F. Ament, and C. Simmer (2008), Validation of TERRA-ML with discharge measurements, *Meteorol. Z.*, *17*(6), 763–773.
- Greve, P., K. Warrach-Sagi, and V. Wulfmeyer (2013), Evaluating soil water content in a WRF-NOAH downscaling experiment, *J. Appl. Meteorol. Climatol.*, *52*, 2312–2327, doi:10.1175/JAMC-D-12-0239.1.
- Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, *Hydrol. Processes*, *22*(18), 3802–3813.
- Gupta, H. V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye (2012), Towards a comprehensive assessment of model structural adequacy, *Water Resour. Res.*, *48*, W08301, doi:10.1029/2011WR011044.
- Hemming, D., R. Betts, and M. Collins (2013), Sensitivity and uncertainty of modelled terrestrial net primary productivity to doubled CO₂ and associated climate change for a relatively large perturbed physics ensemble, *Agric. For. Meteorol.*, *170*, 79–88.
- Höglind, M., S. M. Thorsen, and M. A. Semenov (2013), Assessing uncertainties in impact of climate change on grass production in northern Europe using ensembles of global climate models, *Agric. For. Meteorol.*, *170*, 103–113.
- Hohenegger, C., P. Brockhaus, C. Bretherton, and C. Schar (2009), The soil moisture-precipitation feedback in simulations with explicit and parameterized convection, *J. Clim.*, *22*(19), 5003–5020.
- Ingwersen, J., et al. (2011), Comparison of Noah simulations with eddy covariance and soil water measurements at a winter wheat stand, *Agric. For. Meteorol.*, *151*(3), 345–355.
- IUSS Working Group WRB (2007), World reference base for soil resources 2006, first update 2007, *World Soil Resour. Rep.* 103, Food Agric. Organ., Rome.
- Jarvis, P. G. (1976), The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field, *Philos. Trans. R. Soc. London B*, *273*, 593–610.
- Koster, R. D., et al. (2006), Glace: The global land-atmosphere coupling experiment. Part I: Overview, *J. Hydrometeorol.*, *7*(4), 590–610, doi:10.1175/JHM510.1.
- Kothavala, Z., M. A. Arain, T. A. Black, and D. Verseghy (2005), The simulation of energy, water vapor and carbon dioxide fluxes over common crops by the Canadian land surface scheme (CLASS), *Agric. For. Meteorol.*, *133*, 89–108.
- Krinner, G., N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cycles*, *19*, GB1015, doi:10.1029/2003GB002199.
- Lohmann, D., E. Raschke, B. Nijssen, and D. Lettenmaier (1998), Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model, *Hydrol. Sci. J.*, *43*, 131–142.
- Mahecha, M. D., et al. (2010), Comparing observations and process-based simulations of biosphere-atmosphere exchanges on multiple timescales, *J. Geophys. Res.*, *115*, G02003, doi:10.1029/2009JG001016.
- Maxwell, R. M., and N. L. Miller (2005), Development of a coupled land surface and groundwater model, *J. Hydrometeorol.*, *6*(3), 233–247, doi:10.1175/JHM422.1.
- Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, *10*(3), 282–290.
- Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *J. Geophys. Res.*, *110*, D21106, doi:10.1029/2005JD006111.
- Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with gravity recovery and climate experiment data, *J. Geophys. Res.*, *112*, D07103, doi:10.1029/2006JD007522.
- Niu, G. Y., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, *116*, D12109, doi:10.1029/2010JD015139.
- Noblet-Ducoudré, N. d., S. Gervois, P. Ciais, N. Viovy, N. Brisson, B. Seguin, and A. Perrier (2004), Coupling the soil-vegetation-atmosphere-transfer scheme ORCHIDEE to the agronomy model STICS to study the influence of croplands on the European carbon and water budgets, *Agronomie*, *24*, 397–407.
- Oleson, K., et al. (2004), Technical description of the community land model (CLM), *NCAR Tech. Note NCAR/TN-461+STR*, Natl. Cent. Atmos. Res., Boulder, Colo.
- Oleson, K. W., et al. (2008), Improvements to the community land model and their impact on the hydrological cycle, *J. Geophys. Res.*, *113*, G01021, doi:10.1029/2007JG000563.
- Overgaard, J., D. Rosbjerg, and M. Butts (2006), Land-surface modelling in hydrological perspective—A review, *Biogeosciences*, *3*, 229–241.
- Rosenzweig, C., et al. (2013), The agricultural model intercomparison and improvement project (AgMIP): Protocols and pilot studies, *Agric. For. Meteorol.*, *170*, 166–182.
- Rosero, E., Z.-L. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G.-Y. Niu (2010), Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, *J. Geophys. Res.*, *115*, D03106, doi:10.1029/2009JD012035.
- Rosero, E., L. E. Gulden, Z.-L. Yang, L. G. De Goncalves, G.-Y. Niu, and Y. H. Kaheil (2011), Ensemble evaluation of hydrologically enhanced Noah-LSM: Partitioning of the water balance in high-resolution simulations over the Little Washita River experimental watershed, *J. Hydrometeorol.*, *12*(1), 45–64, doi:10.1175/2010JHM1228.1.
- Rötter, R. P., T. R. Carter, J. E. Olesen, and J. R. Porter (2011), Crop-climate models need an overhaul, *Nat. Clim. Change*, *1*(4), 175–177, doi:10.1038/nclimate1152.
- Santanello, J. A., C. D. Peters-Lidard, and S. V. Kumar (2011), Diagnosing the sensitivity of local land-atmosphere coupling via the soil moisture-boundary layer interaction, *J. Hydrometeorol.*, *12*(5), 766–786, doi:10.1175/JHM-D-10-05014.1.
- Schaake, J. C., V. I. Koren, Q.-Y. Duan, K. Mitchell, and F. Chen (1996), Simple water balance model for estimating runoff at different spatial and temporal scales, *J. Geophys. Res.*, *101*(D3), 7461–7475.

- Schädler, G. (2007), A comparison of continuous soil moisture simulations using different soil hydraulic parameterizations for a site in Germany, *J. Appl. Meteorol. Climatol.*, *46*, 1275–1289.
- Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher (1986), A simple biosphere model (SIB) for use within general circulation models, *J. Atmos. Sci.*, *43*(6), 505–531.
- Wang, E. (1997), *Development of a Generic Process-Oriented Model for Simulation of Crop Growth*, Herbert Utz Verlag Wissenschaft, München.
- Wang, E., and T. Engel (2000), SPASS: A generic process-oriented crop model with versatile windows interfaces, *Environ. Modell. Software*, *15*, 179–188.
- Warrach-Sagi, K., V. Wulfmeyer, R. Grasselt, F. Ament, and C. Simmer (2008), Streamflow simulations reveal the impact of the soil parameterization, *Meteorol. Z.*, *17*(6), 751–762.
- Wöhling, T., et al. (2013), Multiresponse, multiobjective calibration as a diagnostic tool to compare accuracy and structural limitations of five coupled soil-plant models and CLM3.5, *Water Resour. Res.*, *49*, 8200–8221, doi:10.1002/2013WR014536.
- Wolf, A. (2011), Estimating the potential impact of vegetation on the water cycle requires accurate soil water parameter estimation, *Ecol. Modell.*, *222*(15), 2595–2605.
- Wood, E. F., et al. (1998), The project for intercomparison of land-surface parameterization schemes (PILPS) phase 2(c) Red-Arkansas River basin experiment: 1. Experiment description and summary intercomparisons, *Global Planet. Change*, *19*, 115–135.
- Xue, Y., P. J. Sellers, J. L. Kinter, and J. Shukla (1991), A simplified biosphere model for global climate studies, *J. Clim.*, *4*(3), 345–364.
- Yang, Z.-L., and R. E. Dickinson (1996), Description of the biosphere-atmosphere transfer scheme (BATS) for the soil moisture workshop and evaluation of its performance, *Global Planet. Change*, *13*, 117–134.
- Yang, Z. L., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, *J. Geophys. Res.*, *116*, D12110, doi:10.1029/2010JD015140.
- Zhou, X., Y. Zhang, Y. Wang, H. Zhang, J. Vaze, L. Zhang, Y. Yang, and Y. Zhou (2012), Benchmarking global land surface models against the observed mean annual runoff from 150 large basins, *J. Hydrol.*, *470–471*, 269–279.