

SPECIAL SECTION: TRANSDISCIPLINARY CONTRIBUTIONS AND OPPORTUNITIES IN SOIL PHYSICAL HYDROLOGY

Crop growth and soil water fluxes at erosion-affected arable sites: Using weighing lysimeter data for model intercomparison

Jannis Groh^{1,2}  | Efstathios Diamantopoulos³  | Xiaohong Duan⁴ | Frank Ewert¹ | Michael Herbst²  | Maja Holbak³  | Bahareh Kamali¹  | Kurt-Christian Kersebaum^{1,5}  | Matthias Kuhnert⁶ | Gunnar Lischeid¹  | Claas Nendel¹  | Eckart Priesack⁴  | Jörg Steidl¹  | Michael Sommer¹ | Thomas Pütz²  | Harry Vereecken²  | Evelyn Wallor¹ | Tobias K.D. Weber⁷  | Martin Wegehenkel¹ | Lutz Weihermüller²  | Horst H. Gerke¹ 

¹ Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany

² Forschungszentrum Jülich, Agrosphere, Institute of Bio- and Geoscience IBG-3, Jülich, Germany

³ Dep. of Plant and Environmental Science, Univ. of Copenhagen, Copenhagen, Denmark

⁴ Helmholtz Zentrum München-German Research Center for Environmental Health, Neuherberg, Germany

⁵ Global Change Research Institute, Czech Academy of Sciences, Brno, Czech Republic

⁶ Institute of Biological and Environmental Science, Univ. of Aberdeen, Aberdeen, UK

⁷ Institute of Soil Science and Land Evaluation, Univ. of Hohenheim, Stuttgart, Germany

Abstract

Agroecosystem models need to reliably simulate all biophysical processes that control crop growth, particularly the soil water fluxes and nutrient dynamics. As a result of the erosion history, truncated and colluvial soil profiles coexist in arable fields. The erosion-affected field-scale soil spatial heterogeneity may limit agroecosystem model predictions. The objective was to identify the variation in the importance of soil properties and soil profile modifications in agroecosystem models for both agronomic and environmental performance. Four lysimeters with different soil types were used that cover the range of soil variability in an erosion-affected hummocky agricultural landscape. Twelve models were calibrated on crop phenological stages, and model performance was tested against observed grain yield, aboveground biomass, leaf area index, actual evapotranspiration, drainage, and soil water content. Despite considering identical input data, the predictive capability among models was highly diverse. Neither a single crop model nor the multi-model mean was able to capture the observed differences between the four soil profiles in agronomic and environmental

Abbreviations: AgBio, total aboveground biomass; AGC, AgroC-model; BBCH, phenological stages; Dd_1, Calcic Luvisol; Dd_5, Nudiargic Luvisol; DDC, DailyDayCent; DY, Daisy; ENCE, Expert-N-CERES; ENGE, Expert-N-GECCROS; ENSP, Expert-N-SPASS; ENSU, Expert-N-SUCROS; ET, evapotranspiration; Gr_K, Calcic Regosol; GY, grain yield; Hd_S, Colluvic Regosol; HER, HERMES; LAI, leaf area index; MMM, multi-model mean; MON, MONICA; NetQ, net water flux across the lysimeter bottom at 1.5-m soil depth; SWC, soil water content; TERENO, TERrestrial ENvironmental Observatories; TH, THESEUS; THD, THESEUS-Hydrus-1D; THGS, THESEUS-HydroGeoSphere; ZALF, Leibniz Centre for Agricultural Landscape Research

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Vadose Zone Journal* published by Wiley Periodicals LLC on behalf of Soil Science Society of America

Correspondence

Jannis Groh, Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany.
Email: groh@zalf.de, j.groh@fz-juelich.de

[Correction added on September 23, 2020, after first online publication: Projekt Deal funding statement has been added.]

variables. The model's sensitivity to soil-related parameters was apparently limited and dependent on model structure and parameterization. Information on phenology alone seemed insufficient to calibrate crop models. The results demonstrated model-specific differences in the impact of soil variability and suggested that soil matters in predictive agroecosystem models. Soil processes need to receive greater attention in field-scale agroecosystem modeling; high-precision weighable lysimeters can provide valuable data for improving the description of soil–vegetation–atmosphere process in the tested models.

1 | INTRODUCTION

Agroecosystem models are used for an increasing range of applications, including the assessment of risk of climate change on food security and the development of suitable management strategies for adaptation to and mitigation of climate change (Challinor et al., 2018; Martre et al., 2015). A large number of agroecosystem models, differing in model complexity and functionality to simulate and predict the major physiological processes that determine crop growth and yield and their responses to environmental and management factors, have been developed and used for risk assessment studies on agricultural production (Asseng et al., 2013; O'Leary et al., 2015). Each of these models comes with an inherent prediction error when used to predict the response of crop growth and soil water flux dynamics to climatic conditions such as changing seasonal rainfall patterns, rising temperatures, and elevated atmospheric CO₂ concentrations. Therefore, comparative studies using different models are required to evaluate and quantify the uncertainties of simulations results. Several recent studies intercompared crop models and systematically analyzed the sensitivity to different climatic conditions for different crops: wheat (*Triticum aestivum* L.; Asseng et al., 2013, 2014, 2019; Frieler et al., 2017; Hussain, Khaliq, Ahmad, & Akhtar, 2018; Martre et al., 2015; O'Leary et al., 2015; Palosuo et al., 2011; Schaubberger et al., 2017; Wallach et al., 2018), barley (*Hordeum vulgare* L.; Rötter et al., 2012; Salo et al., 2016), rice (*Oryza sativa* L.; Frieler et al., 2017; Hasegawa et al., 2017; Li et al., 2015), maize (*Zea mays* L.; Bassu et al., 2014; Durand et al., 2018; Frieler et al., 2017; Schaubberger et al., 2017), potato (*Solanum tuberosum* L.; Fleisher et al., 2017), soybean [*Glycine max* (L.) Merr.; Battisti, Sentelhas, & Boote, 2017; Battisti et al., 2018; Frieler et al., 2017; Schaubberger et al., 2017], sugarcane (*Saccharum* spp.; Marin, Thorburn, Nassif, & Costa, 2015), and crop rotations (Kollas et al., 2015; Yin, Kersebaum, Kollas, Baby, et al., 2017; Yin, Kersebaum, Kollas, Manevski, et al., 2017). In several of these studies (Kimball et al., 2019; Martre et al., 2015; Palosuo

et al., 2011; Yin, Kersebaum, Kollas, Baby, et al., 2017; Yin, Kersebaum, Kollas, Manevski, et al., 2017), model calibration was mostly based on experimental data with the focus on agronomic in-season variables such as leaf area index (LAI), phenological stages (BBCH), and end-of-season variables such as plant N content, grain yield (GY), or total aboveground biomass (AgBio) at harvest. The explicit consideration of soil water flux dynamics in crop growth intercomparison studies has been largely neglected despite the importance of water and nutrient availability for crop growth in field conditions and simulations. Regardless of its relevance to increase the reliability and confidence of such modeling tools, few simulation studies have characterized the implications of using, in addition to agronomic variables such as crop growth, environmental- and ecosystem-related variables such as actual evapotranspiration (ET_a; Cammarano et al., 2016; Kimball et al., 2019; Wöhling, Schöniger, Gayler, & Nowak, 2015), spatial soil textural information (Hoffmann et al., 2016; Mahajan et al., 2019), subsoil N content (Wallor et al., 2018), and soil water content (SWC; Wallor et al., 2018; Wöhling et al., 2013, 2015). To accurately describe and understand the crop and soil agroecosystem processes, soil water (and matter) flux-related processes should also be part of the model calibration strategy, because these variables (i.e., ET_a, SWC, storage, and drainage) are critical, as they influence the crop model predictions. Unfortunately, most model calibrations including those water and matter fluxes are performed with only a few temporally or spatially resolved data (Martre et al., 2015; Žydelis, Weihermüller, Herbst, Klosterhalfen, & Lazauskas, 2018). Thus, classical agroecosystem model intercomparison studies focused their model calibration on crop-relevant data only, rather than on soil water dynamics (Seidel, Palosuo, Thorburn, & Wallach, 2018). This lack in the calibration is mostly because datasets that include all necessary information are often not available (Kersebaum et al., 2015). Considering a broader range of data types in the crop model calibration can lead to a more accurate description of agroecosystem or land surface models (Wöhling et al., 2013). It

should be noted that the way models are calibrated may also be important for the reliability of the models predictions. However, a recent investigation using ET_a in crop model intercomparison studies showed that much of the difference and variability in predicting ET_a by crop models was mainly a consequence of using different approaches to estimate potential ET (Kimball et al., 2019). This example demonstrates that in addition to differences in crop model complexity and model functionality, the selection of boundary conditions (Diamantopoulos et al., 2017) can largely affect the outcome of crop models. Thus, not only the selection of the upper boundary (e.g., reference ET) but also the selection of the bottom boundary condition might be relevant in model intercomparison studies if water from deeper soil layers or the water table is available for capillary rise and may serve as an additional water supply for root water uptake and evaporation (Groh, Vanderborght, Pütz, & Vereecken, 2016; Luo & Sophocleous, 2010). The bottom boundary condition can also indirectly affect the water uptake from aboveground sources (irrigation and precipitation), because the plant consumption of infiltrating water differs in the case of upward-directed flow in the subsoil layer, and this eventually affects the crop model predictions. This highlights the need to use standardized model boundary conditions in model intercomparison studies. We therefore hypothesize that agronomic and environmental observations obtained with cropped weighing lysimeters are suitable for testing agroecosystem models, as they provide information about all water and matter fluxes, and information about boundary conditions and crop performance. When applying agroecosystem models, the field-scale soil variability needs to be considered. In arable fields, for example, spatial soil heterogeneity can result from long-term soil management in combination with erosion effects. In hummocky soil landscapes, erosion-affected soil profile modifications can lead to truncated soil profiles at exposed and hillcrest positions and to colluvial soils at toeslopes or in local depressions (Figure 1). Both profile truncation and colluviation changes the existence and thickness of soil horizons, and even the soil hydraulic properties of the same soil horizons can substantially differ (Rieckh, Gerke, & Sommer, 2012). Because of the close interactions of soil- and crop-related biophysical processes, these soil modifications may lead to differences in water and matter fluxes and consequently to differences in crop performance (Gerke, Rieckh, & Sommer, 2016). Quantitative studies on the limitations of model predictions of field-scale agroecosystem simulations caused by different soil characteristics, especially for truncated and colluviated profiles, are rare, even if such phenomena can be detected frequently. The classical representation of arable fields by using dominant soil profiles may limit predictions of both the agronomic

Core Ideas

- We performed model intercomparison based on lysimeter data of erosion-affected soils.
- Uniform crop model calibration was performed by providing only phenological growth stages.
- We compared individual models predictions with those of the multi-model mean.
- Differences between erosion-affected soils could not be described well by models and the multi-model mean.
- We used lysimeter data for testing how soil matters in agroecosystem models.

crop performance as well as the environmental fluxes in erosion-affected soil landscapes (Rieckh, Gerke, Siemens, & Sommer, 2014).

The objective of this crop model comparison study was to quantify differences in the role of soil properties in agroecosystem modeling when trying to predict both agronomic aspects such as crop development and yield and environmental aspects such as water and matter fluxes. Specific aims were (a) to compare forward simulation results after minimal calibration on phenology of crop growth and soil water flux-related ecosystem variables based on lysimeter observations from erosion-affected soils, and (b) to evaluate how well the models reproduce the agronomic and environmental variables of erosion-affected soil profiles.

Data are provided from the TERrestrial ENvironmental Observatories (TERENO)-SOILCan lysimeter network site Dedelow, which is located in the hummocky ground moraine landscape of northeastern Germany (Heinrich et al., 2018; Pütz et al., 2016). This landscape type is important for global crop production, as it covers ~10% of the arable land at the global temperate climatic zone (Chapman et al., 2020; Sommer, Fiedler, Glatzel, & Kleber, 2004). It reveals a regular, erosion- and deposition-related soil pattern across continents (Pennock, Bedard-Haughn, Kiss, & van der Kamp, 2014; Sommer, Gerke, & Deumlich, 2008) and has been intensively studied under aspects of erosion feedbacks on crop biomasses and yields (Papiernik et al., 2005), as well as on greenhouse gas fluxes (Hoffmann et al., 2017; Pennock et al., 2010). We selected lysimeter observations from four soil profiles, which differed in both soil type and soil erosion history and capture a broad range of the field-scale spatial variability of soils in erosion-affected, hummocky agricultural landscapes (Miller et al., 2016; Sommer et al., 2008). For each soil monolith, data were provided for (a) typical agronomic observations (GY,

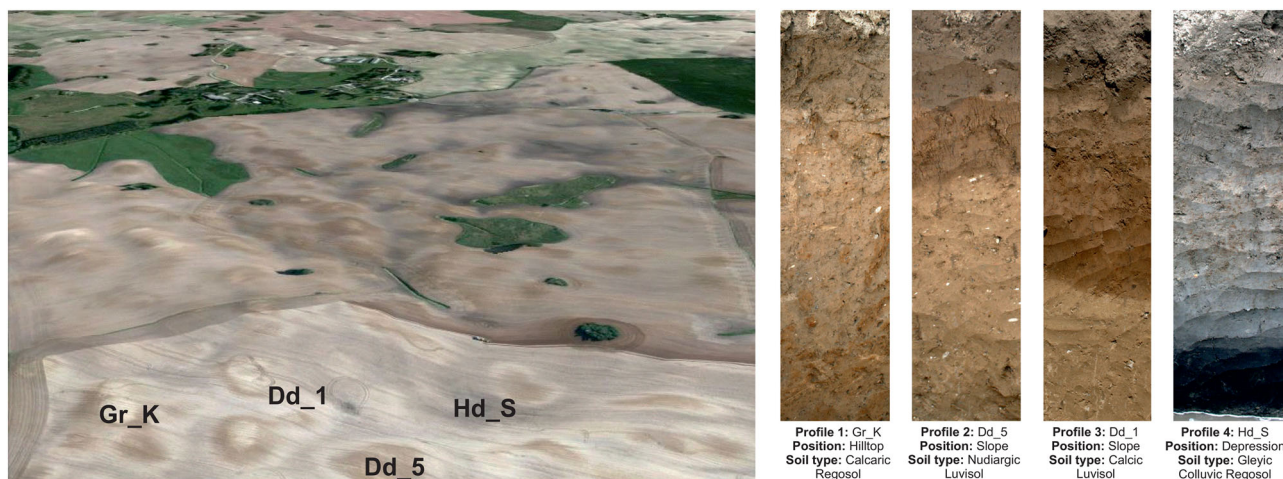


FIGURE 1 Hummocky ground moraine landscape of the Uckermark, northeast Germany (left), and idealized catena representing soil profiles of the lysimeters (right)

AgBio, LAI, and phenology [BBCH]), and (b) environmental ecosystem fluxes and states—namely, ET_a , drainage, and SWC—from August 2014 until October 2018.

Ten agroecosystem models and two crop-soil hydrology models were tested with limited calibration, based on phenology stages only, to predict crop growth and water flux dynamics. The dataset for the simulations included weather, reference ET, soil water retention characteristics, soil initial conditions, site management, targeted range of GY, and BBCH.

2 | MATERIALS AND METHODS

2.1 | Site and soil descriptions

The data for this study were collected at the experimental field station in Dedelow, Germany (53°22'2.45" N, 13°48'10.91" W, 50–60 m asl) and are maintained by the Leibniz Centre for Agricultural Landscape Research (ZALF). The station is located in a hummocky arable soil landscape and is part of the Northeast German Lowland Observatory of TERENO (Heinrich et al., 2018) and the German wide lysimeter network SOILCan (Pütz et al., 2016). In total, 12 high-precision weighing lysimeters (METER Group), each with a surface area of 1 m² and a depth of 1.5 m, were installed at the experimental field. Four characteristic soil types, which evolved from glacial till, were selected for this study. They represent a typical catena of hummocky ground moraines (Figure 1) showing an extremely eroded Calcaric Regosol (Gr_K) at the exposed hilltop, a strongly eroded Luvisol (Nudiargic Luvisol Dd_5) and a noneroded Luvisol (Calcic Luvisol Dd_1) from slopes, and a depositional Colluvic Regosol (Hd_S) at the depression. Physical and chemical soil characteristics

can be taken from Supplemental Tables S1 and S2 and were provided to each modeling group. The soil water retention characteristics were determined from undisturbed soil cores with the evaporation method using the HYPROP technique (METER Group). This method provides drying soil water retention data, which were fitted to the standard van Genuchten-type soil water retention model. Soil texture and soil chemical properties were obtained from disturbed samples. More information on the site and soil properties and the lysimeter setup of SOILCan can be found in Herbrich and Gerke (2017), Herbrich, Gerke, Bens, and Sommer (2017), and Pütz et al. (2016).

The arable-land lysimeters were surrounded by a control field (~2.3 ha). The plant management (crop type, fertilizer, plant protection application, crop growth regulators, and tillage) on the lysimeters and the surrounding field was identical during the observation period (1 Aug. 2014 until 31 Oct. 2018). Soil tillage was carried out manually each year (depth of tillage = 8–12 cm), and fertilizer applications during the observation period are listed in Supplemental Table S3, which was also provided to each modeling group. The lysimeters were cultivated with winter wheat (*T. aestivum*) cultivar 'Julius', winter wheat (*T. aestivum*) cultivar 'Pionier', winter rye (*Secale cereale* L.) cultivar 'SU Cossani', winter rye (*S. cereale*) cultivar 'SU Cossani', and spring oats (*Avena sativa* L.) cultivar 'Ivory'. It should be noted that the winter rye was not harvested in 2018, and the entire plant material was incorporated into the soil by plowing before sowing oats on 11 Apr. 2018. Seasonal crop development was characterized by BBCH and LAI (m² m⁻²) measured by a plant canopy analyzer probe (SunScan Probe Type SS1, Delta-T Devices). Dry matter AgBio and GY were gravimetrically determined at harvest after drying at 60 °C for >24 h after reaching constant weight.

2.2 | Lysimeter data

Observations on water fluxes (ET_a , precipitation, and drainage) were obtained for each soil profile from the corresponding lysimeter mass data. The weighing precision of the lysimeters was 10 g (~ 0.01 mm); for the water reservoir tank used to capture drainage, a precision of 1 g (~ 0.001 mm) was specified. For the Calcaric Regosol (Gr_K) and the Colluvic Regosol (Hd_S), the field pressure head values were obtained from tensiometers installed at the original landscape position of the soil monolith extraction in order to more realistically represent ET processes under conditions of upward-directed soil water fluxes (Groh et al., 2016). At the bottom of the lysimeters (i.e., 1.4- to 1.5-m depth), a porous suction rake device was installed and connected with a pumping system that allowed the bidirectional transfer of water between lysimeter soil and the weighable water reservoir. The pumping system operated according to an algorithm that controlled pumping direction and duration according to differences in pressure heads observed in the lysimeter at 1.4-m depth and those measured in same soil depth at the monolith extraction site's landscape position. The lysimeter weights were logged every minute and post-processed using the AWAT noise filter of Peters et al. (2017). For this study, water fluxes obtained from lysimeter and water reservoir tank mass data at 1-min resolution were aggregated to daily sums of ET_a , precipitation, and net water flux across the lysimeter bottom at 1.5-m soil depth (NetQ).

Lysimeters were equipped with time domain reflectometry (TDR, CS610, Campbell Scientific) sensors to determine SWC at 10-, 30-, and 50-cm soil depths. Average daily SWC of the first 60 cm was calculated by SWC observations from 10-, 30-, and 50-cm depths by using a depth interval weighed sum for the first 60 cm.

Standard meteorological data of air temperature, global radiation, relative humidity, air pressure, and wind speed were provided in 10-min intervals by a weather station (Type WXT510, Vaisala Oyj) that was located ~ 4 m from the lysimeters in Dedelow. A linear regression model between the corresponding meteorological variables at the lysimeter station and surrounding weather stations (Dedelow, Heydenhof, Christianhof, Kuntzow, and Muessentin), obtained from the publicly accessible TERENO platform (<http://www.tereno.net/ddp/>), was used for gap filling of missing data in the time series. We applied a stepwise gap-filling approach, so that missing meteorological values were filled based on the observation that had the highest correlation (adjusted R^2). The gap-filled meteorological parameters were used to estimate reference ET for grass (ET_0) for hourly time steps according to the FAO56 Penman–Monteith method (Allen, Pereira, Raes, & Smith, 1998) and was summed up to daily values. The

parameterization of the bulk surface resistance (soil and plant) and aerodynamic resistance (plant) values in the Penman–Monteith model followed recommendations of Allen et al. (2006). The maximum rooting depth for each crop and the corresponding soil was set in a first step arbitrarily to the depth of the lysimeter (1.5 m). In a second step, the maximum rooting depth was determined based on soil profile observations, assuming that roots penetrate the soil solum down to the C horizon (i.e., highly compacted glacial till) plus the upper 0.2 m of the C horizon along cracks and earthworm burrows, resulting in maximum rooting depths of 0.47 m for Gr_K, 0.85 m for Dd_5, 1.0 m for Dd_1, and 1.50 m for Hd_S.

2.3 | Crop models

This intercomparison study considered 12 simulation models, which have been described in detail in separate publications: AgroC (AGC; Klosterhalfen et al., 2017), DailyDayCent (DDC; Del Grosso et al., 2001, 2006; Parton et al., 2001; Yeluripati et al., 2009); Daisy (DY; Abrahamson & Hansen, 2000; Hansen, Abrahamson, Petersen, & Styczen, 2012); Expert-N (EN; Biernath et al., 2011; Engel & Priesack, 1993; Priesack, Gayler, & Hartmann, 2006) coupled to CERES (ENCE), GECROS (ENGE), SPASS (ENSP), and SUCROS (ENSU); HERMES (HER; Kersebaum, 1995, 2007); MONICA (MON; Nendel et al., 2011; Nendel, Kersebaum, Mirschel, & Wenkel, 2014); THESEUS (TH; Wegehenkel, Luzi, Sowa, Barkusky, & Mirschel, 2019) and two hydrological models, namely, Hydrus-1D (THD; Šimůnek, van Genuchten, & Šejna, 2016) and HydroGeoSphere (THGS; Aquanty, 2013; Brunner & Simmons, 2012). Model simulations were performed on a daily time step, except for the DY and THGS model, which simulated crop growth (DY) and/or water flow (THGS) on an hourly basis. The models represent different degrees of crop model complexity and model functionality to simulate agronomic- and environmental ecosystem fluxes and states. Four crop models shared the same modular structure as EN to simulate soil water flow, soil heat, and N transport but differed in structure to simulate crop growth. The simulation programs THD and THGS used crop outcomes (LAI and root depth) from TH and thus did not have a direct feedback between crop and soil status. However, those models were included in the study to serve as independent reference for Richards equation-based numerical soil hydraulic models as compared with crop models coupled with bucket-type and Richards equation-based soil water models. All other crop models do not share a common system with respect to model structure to simulate crop growth and water flux dynamics.

2.4 | Setup up of model intercomparisons

Each individual modeling group received basic key soil parameters and variables (physical and chemical soil characteristics; Supplemental Tables S1 and S2), daily meteorological conditions (air temperature [maximum, minimum, and mean], wind speed, global radiation, relative humidity, and air pressure), ET_0 , precipitation, bottom boundary condition (distance to groundwater table or pressure head), maximum rooting depth for all crops (fixed and depth adjusted), information on the crop–soil management (sowing density, sowing, and harvesting date), and observed crop phenology (BBCH; Meier, 2001). Hereby, BBCH was provided as the only opportunity allowing for a minimal crop model calibration (Supplemental Table S3). Conversion from simulated development stages to BBCH followed the table of Wang and Engel (1998), except for the models DY and DDC. In DDC, the management events are defined in the input files. The different stages are simulated by a phenological model with crop specific parameters, which were provided in the crop datasets. Information on the range of regional GY was provided for winter wheat (range: 6–12 t ha⁻¹), winter rye (range: 5–11 t ha⁻¹), and oat (range: 2–5 t ha⁻¹) based on long-term yield observations carried out in the ZALF project “AgroScapeLab Quillow” (http://www.zalf.de/en/forschung_lehre).

The initial SWCs supplied to the modeling groups were calculated from the vertical matric potential distribution profile obtained by linearly interpolating between measured values at the top (10 cm) and the bottom (140 cm) of the lysimeter at the starting day. The matric potential values were converted to volumetric water contents using the van Genuchten soil water retention model and the parameters of each soil horizon (Supplemental Table S2).

2.5 | Statistical analysis

The evaluation of the individual model and the multi-model mean (MMM) performance was done for the agronomic ecosystem-related in-season and end-of-season variables (LAI, GY, and AgBio) and the main in-season soil environmental ecosystem variables (ET_a , NetQ, and SWC). We considered the corresponding simulation outputs of 10 crop models to calculate MMM of the model simulations and thus exclude the soil hydraulic models because of missing interactive coupling of crop development with soil moisture status (i.e., THD and THGS) as in the other crop models. The R software (R Core Team, 2016) and the function *nrmse* of the package hydroGOF (Zambrano-Bigiarini, 2017) were used to compare observed with simulated variables. The normalized RMSE (nRMSE) was calculated as

$$nRMSE = \frac{100}{SD(Obs)} \sqrt{\sum_{i=1}^N \frac{(Sim_i - Obs_i)^2}{N}} \quad (1)$$

where SD is standard deviation, N is the number of samples, and Sim and Obs are the simulated and observed values. The nRMSE provides a statistical criterion that allows comparison of the overall model performance variables of each model irrespective of the different dimension of the variable (Wallach, 2006). The normalization in nRMSE was done with the standard deviation rather than the mean of the observation variable. This was necessary, because one variable contains positive and negative values of the water fluxes (NetQ) and SD permits accounting for the range of the variability of the observation and thus enables the comparison of the model performance for different variables of a heterogeneous set of observation data (Coucheney et al., 2015). Spearman's correlation coefficient, ρ , was used to investigate the relationship between the nRMSE of the corresponding agronomic and environmental ecosystem variable.

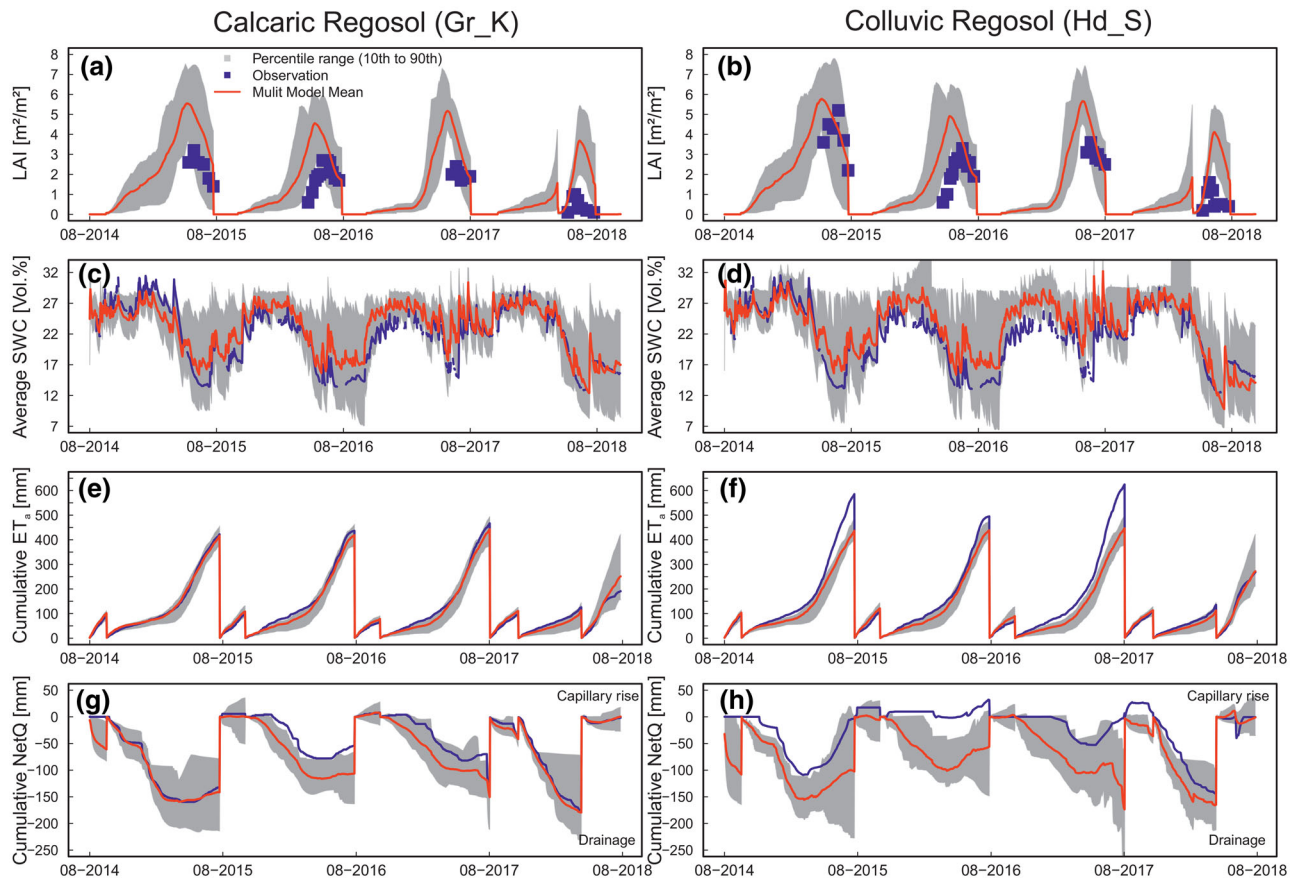
3 | RESULTS AND DISCUSSION

3.1 | Evaluation of agronomic and environmental fluxes and states

The models were calibrated only on phenology data, and model performance values of nRMSE (Table 1) ranged between 17% (model AGC) and 83% (ENGE). However, most models achieved relatively low nRMSE (~30%) and were thus able to describe the observed BBCH well. Note that the variable BBCH was not available from results of DDC, Hydrus-1D, and HydroGeoSphere. For the observation period (August 2014–October 2018), the in-season variable LAI was overestimated by all individual models and the MMM for all four soils (e.g., Figures 2a and 2b). Large overestimation of LAI was especially visible for oats in 2018. Most crop models do not provide well-tested parameter values for spring oats, including their water deficit response, which may be the reason why dry conditions (early ripening) in 2018 resulted in this large deviation from observations. The variability in crop models for in-season variables (LAI, SWC, ET_a , and NetQ) was generally large, and observations were, in part, outside of the 10th to 90th percentile range of the simulations. For example, in season 2016–2017 the cumulative simulated MMM and observed ET_a and NetQ showed a discrepancy of 178 and –185 mm, respectively (see Figures 2f and 2h). However, the MMM captured the temporal dynamics of the

TABLE 1 Normalized RMSE (nRMSE) calculated from the simulated and observed phenology stages (BBCH) for each model

Model	nRMSE			
	Calcaric Regosol (Gr_K)	Nudiargic Luvisol (Dd_5)	Calcic Luvisol (Dd_1)	Colluvic Regosol (Hd_S)
	%			
AgroC	16.6	17.3	17.2	16.6
DailyDayCent	No data	No data	No data	No data
Daisy	19.3	19.1	18.8	19.4
Expert-N-CERES	33.3	33.2	34.5	33.2
Expert-N-GECCROS	82.7	82.3	82.8	82.6
Expert-N-SPASS	21.0	21.5	20.8	20.8
Expert-N-SUCROS	47.0	46.8	45.8	47.0
HERMES	19.8	19.7	19.4	18.9
HydroGeoSphere	No data	No data	No data	No data
Hydrus-1D	No data	No data	No data	No data
MONICA	32.2	31.3	31.0	32.5
Theseus	37.2	37.1	36.1	37.9

**FIGURE 2** Measured and simulated in-season variables (a, b) leaf area index (LAI), (c, d) average soil water content (SWC, 0–0.6 m), (e, f) cumulative seasonal actual evaporation and evapotranspiration (ET_a), and (g, h) net soil water flux at 1.5-m soil depth (NetQ) shown exemplarily for the most contrasting soil profiles (left: Gr_K; right: Hd_S) for four growing seasons over the simulation period (1 Aug. 2014 until 31 Oct. 2018). The gray shaded area marks the 10th to 90th percentile of the values simulated by different crop models. The red line depicts the multi-model mean (MMM), and the blue dots (LAI) or the blue line (SWC, ET_a , and NetQ) depict the observations

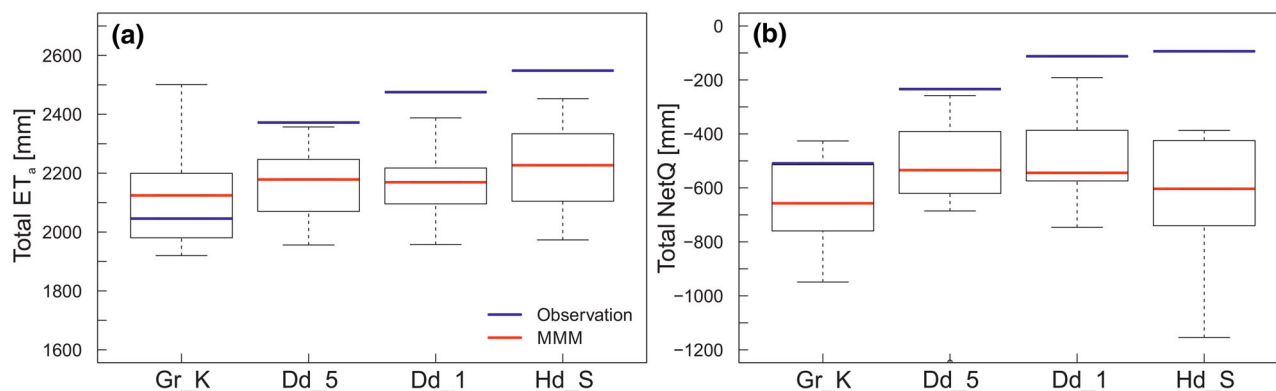


FIGURE 3 Cumulative simulated and observed (blue lines) results of (a) actual evapotranspiration (ET_a) and (b) net water flux at the lysimeter bottom (NetQ) for four soil profiles (Calcaric Regosol [Gr_K], Nudiargic Luvisol [Dd_5], Calcic Luvisol [Dd_1], and Colluvic Regosol [Hd_S]) over the simulation period (1 Aug. 2014 until 31 Oct. 2018); the simulations were obtained from 10 different crop models; the red lines depict the multi-model mean (MMM)

SWC well, despite the large range of simulated SWC of the individual models. Crop models generally underestimated the seasonal ET_a and overestimated the seasonal NetQ (e.g., Figures 2f and 2h).

The overestimation of the LAI, on the one hand, and the underestimation of ET_a , on the other hand, suggest that the partitioning of ET_a into its constituent components, actual evaporation (E_a) and actual transpiration (T_a), and thus the crop water stress conditions in terms of the ratio T_a to potential transpiration (T_p), was perhaps insufficiently described by the crop models. The LAI and ET_a are dependent variables, at least for models HER and MON, which means that a failure in simulating LAI consequently leads to a failure in simulating ET_a . However, for some models ET is not linked to the crop, because T_a is not calculated as a function of LAI directly (e.g., DDC, Hydrus-1D, and HydroGeoSphere). The discrepancy between simulated and observed ET_a is highly relevant when estimating crop water use efficiency, and crop productivity (Cammarano et al., 2016).

The observations showed (see Figure 3a, blue lines) that ET_a increased along the catena from the Calcaric Regosol (Gr_K, extremely eroded) towards the Colluvic Regosol (Hd_S, depositional). The opposite was found for NetQ (Figure 3b, blue lines), where the amount of water that drained from the soil column was larger for the Calcaric Regosol than for the Colluvic Regosol. The dependency of water flux rates on the erosion-affected differences in the soil profiles can be related to the soil water storage capacities, which differ due to erosion or deposition processes. This is in agreement with the results reported by Herbrich, Gerke, Bens, and Sommer (2017), who showed for various Luvisols that the cumulative drainage increased with the degree of erosion-induced soil profile truncation.

The range in measured total ET_a and NetQ between soil profile Gr_K and Hd_S was 2,046 and 2,548 mm for ET_a and -509 and -94 mm for NetQ (negative flux = outflow of the soil profile). In comparison, simulated total ET_a and NetQ ranged only between 2,124 (Gr_K) and 2,227 mm (Hd_S) and between -657 (Gr_K) and -534 mm (Dd_5), respectively.

Long-term soil erosion and deposition changed soil horizon depths (Bt or C nearer to the soil surface in case of erosion) and created new soil horizons (fertile, colluvial soil layers in depositional area). This results in changes of depth functions of soil properties, which are crucial for plant growth (e.g. texture, bulk density, continuity of the pore system, soil organic C, and nutrients). The magnitude of these changes is related to spatially distributed intensities of soil erosion by water and tillage (Gerke & Hierold, 2012), the distance to the ground water table (Rieckh et al., 2012), and their combined effect on the above- and below-ground plant growth and development (Große, 1963; Herbrich, Gerke, & Sommer, 2017; Li, Lobb, Lindstrom, & Farenhorst, 2007). As the simulation results indicate, the average range of the individual crop model simulations and the MMM cannot capture such complex interactions on the variables ET_a and NetQ (see Figures 3a and 3b), when using the provided soil hydraulic properties and calibrating on phenology stages only. However, the field-scale heterogeneity of ET_a and NetQ, which is related to soil spatial variability (e.g., by variations in structure and depth of soil horizons), needs to be considered when simulating crop growth in hummocky landscapes with erosion-affected pedogenesis. Interactions between erosion effects on the water fluxes can further result from heterogeneity of root distribution at specific landscape positions and might also affect solute transport (e.g., dissolved C; Gerke et al., 2016).

TABLE 2 Observed grain yield (GY) for each soil profile for winter wheat, winter rye, and oat and mean GY over the period from 2014 until 2018

Soil profiles		Calcaric Regosol (Gr_K)	Nudiargic Luvisol (Dd_5)	Calcic Luvisol (Dd_1)	Colluvic Regosol (Hd_S)
Year	Crop	t ha ⁻¹			
2014–2015	Winter wheat	6.73	9.03	9.41	12.13
2015–2016	Winter wheat	5.76	7.84	8.24	7.90
2016–2017	Winter rye	5.25	8.79	9.00	10.80
2018	Oats	1.99	2.92	3.88	3.34
2014–2018	Mean ^a	4.93	7.14	7.63	8.54

^aMean grain yield over all three crops (winter wheat, winter rye, and oat).

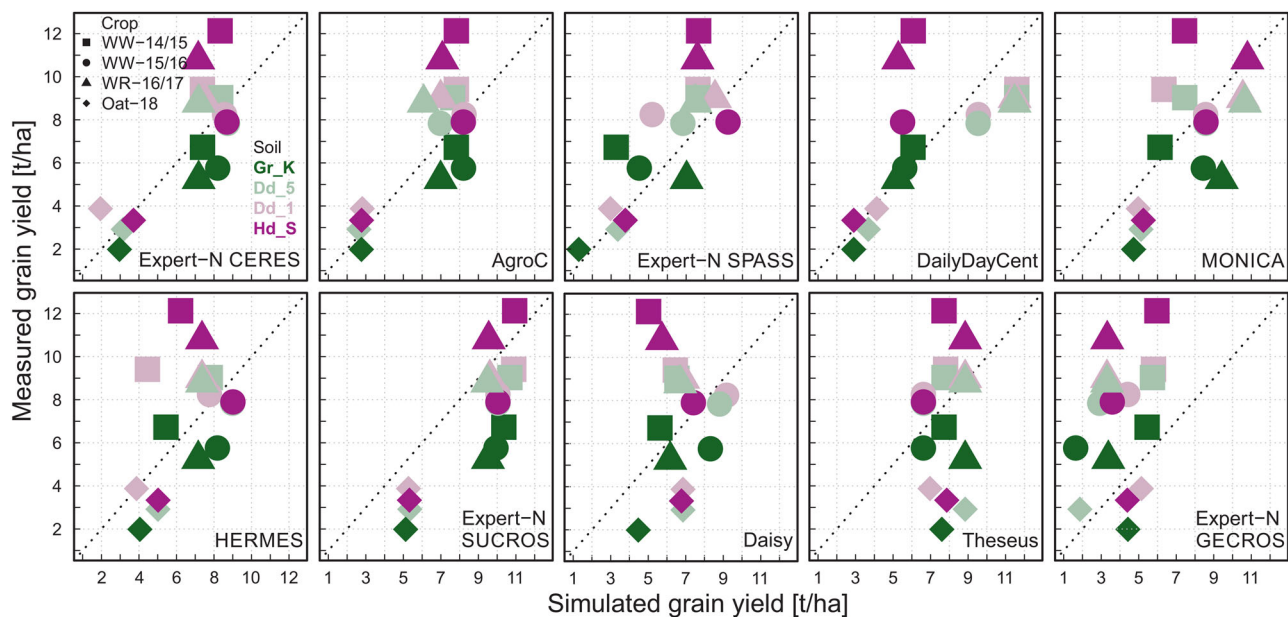


FIGURE 4 Measured and simulate values of the seasonal grain yield (GY) for four different soil profiles (Calcaric Regosol [Gr_K], Nudiargic Luvisol [Dd_5], Calcic Luvisol [Dd_1], and Colluvic Regosol [Hd_S], depicted by different colors). The crop rotation during the observation period from 1 Aug. 2014 until 31 Oct. 2018 was 2014–2015 winter wheat (WW-14/15, square), 2015–2016 winter wheat (WW-15/16, dot), 2016–2017 winter rye (WR-16/17, triangle), and 2018 summer oat (Oat-18; diamond). The simulations were obtained from 10 different crop model simulations

A similar effect of erosion-affected soil profiles could be observed for the measured variable GY. The mean GY, which was calculated over all three crops (winter wheat, winter barley, and oat) from the observation period (2014–2018), clearly shows an increase of mean GY from the Calcaric Regosol (Gr_K, 4.9 t ha⁻¹) to the Colluvic Regosol (Hd_S, 8.5 t ha⁻¹, Table 2). A similar effect could be observed for individual vegetation periods of the winter crops in 2014–2015 and 2016–2017, which showed, in general, the highest GY for the depositional soil Hd_S (Table 2) and the smallest GY for the most eroded soil (Gr_K) along the idealized catena. This was also observed for seasonal measured maximum LAI (not shown). Earlier simulations by Gerke et al. (2016) showed that truncated soils at the hilltop achieved a reduced crop growth and an increased

percolation in comparison to other soils. For winter wheat (2015–2016) observations on GY of the Colluvic Regosol (Hd_S, 7.9 t ha⁻¹), Calcic Luvisol (Dd_1, 8.2 t ha⁻¹), and Nudiargic Luvisol (Dd_5, 7.8 t ha⁻¹) achieved similar values but exceeded the GY of the Gr_K (5.8 t ha⁻¹) by far. The observed range of GY for winter wheat and winter rye agreed well with the range of observed GY from the region (Raatz et al., 2019; Schils et al., 2018; Verch, Kächele, Hörtl, Richter, & Fuchs, 2009). The GY for oats in the dry year, 2018, was highest for the Calcic Luvisol (Dd_1, 3.9 t ha⁻¹), followed by the Colluvic Regosol (Hd_S, 3.3 t ha⁻¹), and the Nudiargic Luvisol (Dd_5, 2.9 t ha⁻¹). The Calcaric Regosol showed also for oats the lowest GY (Gr_K, 2.0 t ha⁻¹). Water stress during high temperature periods in June and July caused an earlier ripening of oats in 2018 (Dr. Verch,

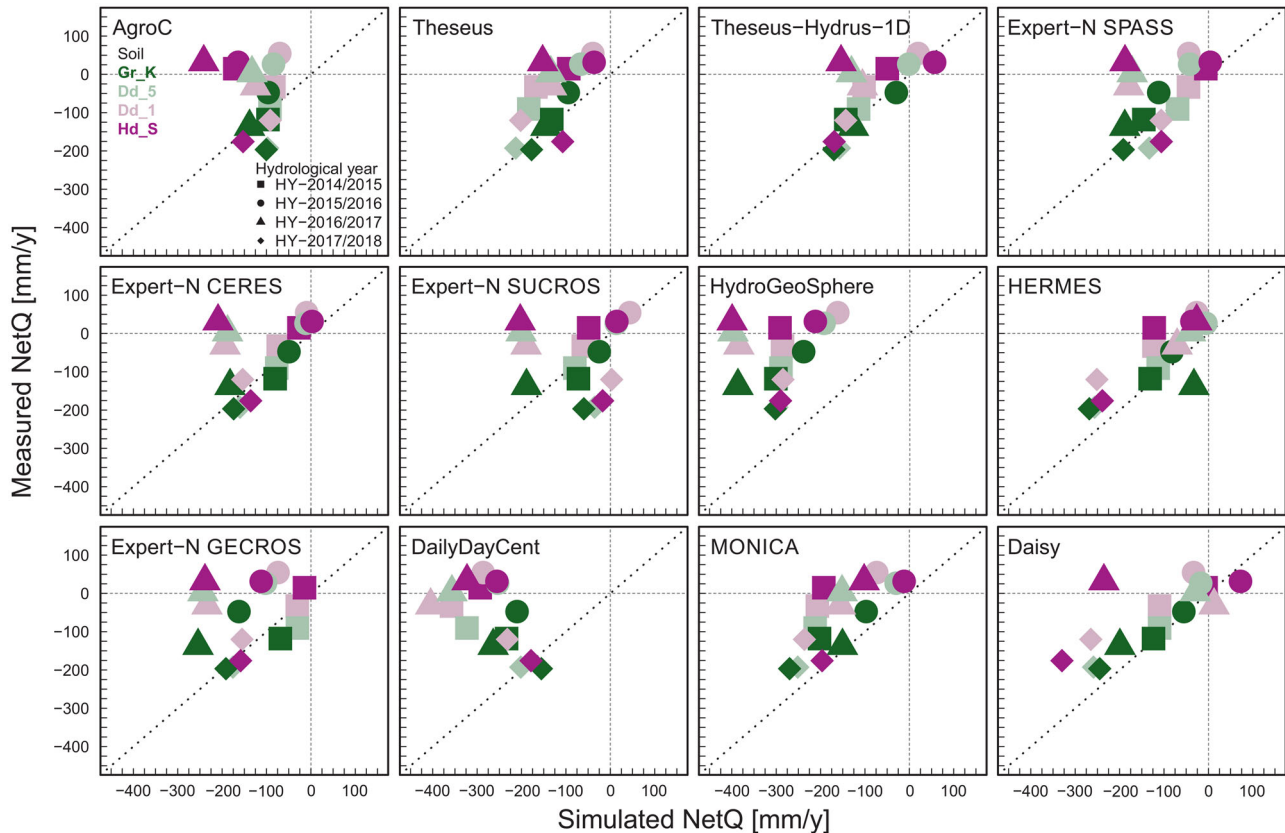


FIGURE 5 Measured and simulated values of the seasonal cumulative net water flux across the lysimeter bottom (NetQ) for four different soil profiles (Calcaric Regosol [Gr_K, dark green], Nudiargic Luvisol [Dd_5, light green], Calcic Luvisol [Dd_1, pink], and Colluvic Regosol [Hd_S, violet]) and crops (solid symbols). The crop rotation during the observation period from 1 Aug. 2014 until 31 Oct. 2018 was winter wheat (2014–2015, square), winter wheat (2015–2016, circle), winter rye (2016–2017, triangle), and summer oat (2018, diamond). The simulations were obtained from 10 different crop and two hydrological model simulations

ZALF Research Station Dedelow, personal communication, 2019), and heat stress during anthesis decreased GY. This agrees well with results from Rezaei et al. (2018), who showed that a combined occurrence of heat and drought stress can lead to a significant decrease in GY.

The deviation of the simulated GY from the 1:1 line in Figure 4 clearly shows that most crop models could not capture how GYs were affected by different erosion-affected soils, as simulated differences in GY between the soil profiles and crops were relatively small for most models. Some of the crop models underestimated the impact of drought stress on simulated yields. For example, TH predicted hardly any crop water stress conditions in some years in terms of the ratio T_a to T_p (not shown here) for all soil profiles. The representation of these interactions between soil- and crop-related biophysical processes needs to be considered when applying agroecosystem models to landscapes with a high soil-spatial variability.

Differences between simulations and observations were also large for seasonal cumulative NetQ (hydrological years, see Figure 5). Most models had problems capturing observed seasonal cumulative NetQ. However, devi-

ations between simulated and observed drainage are not surprising, as crop models were only calibrated on phenology and have rarely been tested with drainage data, simply because of a lack of data from most field experiments (Kersebaum et al., 2015). The use of calibrated instead of measured soil hydraulic properties (Žydelis et al., 2018) could have reduced the deviation of the simulated from the measured NetQ (Figure 5). This approach was not pursued here because a model-specific calibration of soil hydraulic parameters would have limited the comparability between the model results. However, we note that the method of obtaining knowledge about the soil hydraulic properties using the evaporation method might be sensitive to the simulated water fluxes for the following reasons: the intact soil cores used for this method represent only a relatively small volume of the soil horizon; soil hydraulic parameters were determined from the main drying curve of the soil water retention function, and thus dynamic changes of the soil pore structure were due to swelling and shrinkage or wetting and drying; and topsoil tillage operations were neglected. Earlier analyses of data from the same lysimeters showed by comparing laboratory-measured soil

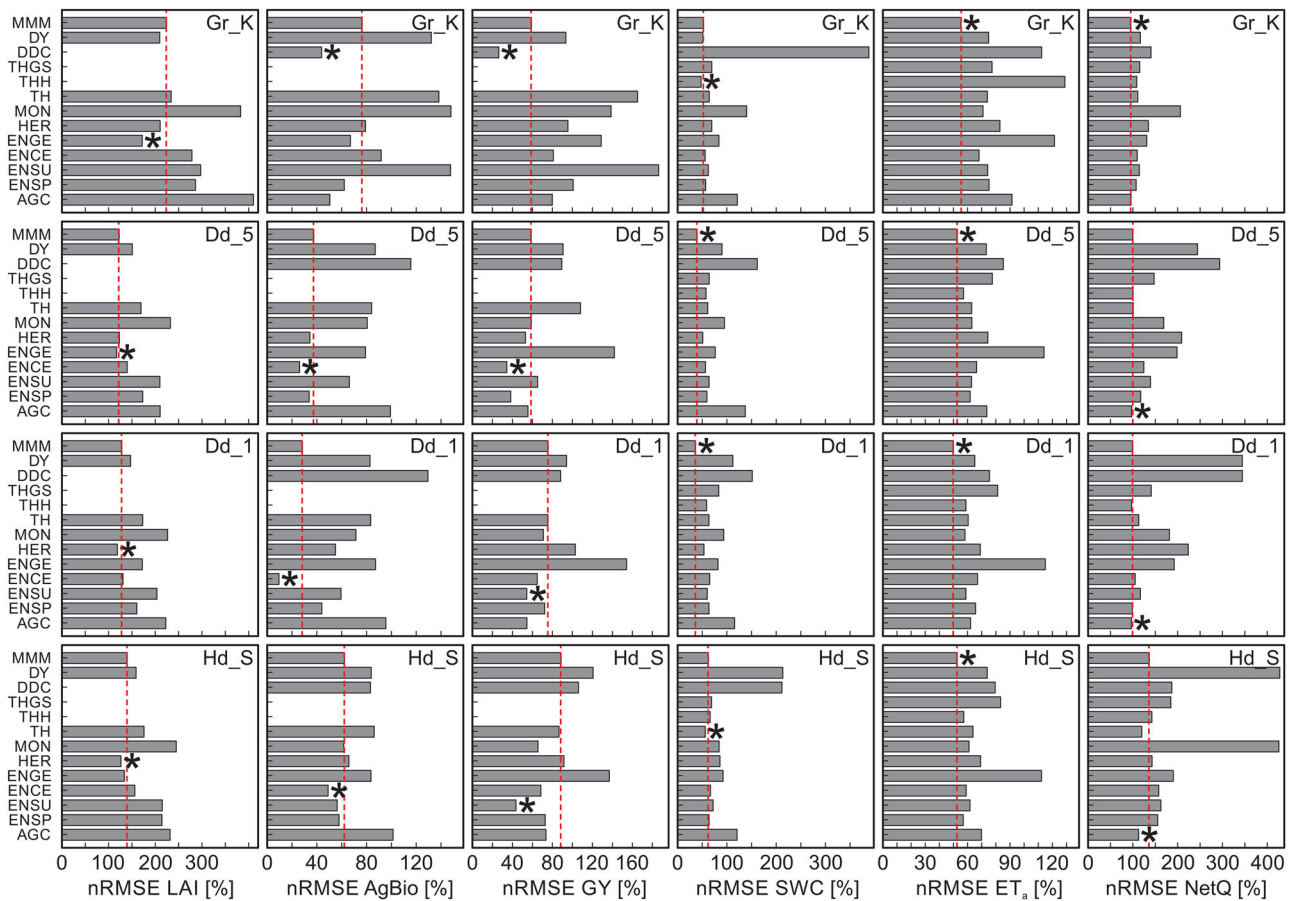


FIGURE 6 Normalized RMSE (nRMSE) between the simulated and observed agronomic and environmental ecosystem variables for the four soils (Calcaric Regosol [Gr_K], Nudiargic Luvisol [Dd_5], Calcic Luvisol [Dd_1], and Colluvic Regosol [Hd_S]), individual model results, and multi-model mean (MMM). The variables include four in-season variables (daily) of leaf area index (LAI), soil water content of the first 0.6 m (SWC), actual evapotranspiration (ET_a), net water flux across the lysimeter bottom at 1.5-m soil depth (NetQ), and two end-of-season values of aboveground biomass (AgBio) and grain yield (GY). The asterisk marks the simulation approach with the best model performance of the agroecosystem variables, and the dashed line (red) indicates the results for the MMM. Model abbreviations: DY, Daisy; DDC, DailyDayCent; THGS, THESEUS-HydroGeoSphere; THH, THESEUS-Hydrus-1D; TH, THESEUS; MON, MONICA; HER, HERMES; ENGE, Expert-N-GECROS; ENCE, Expert-N-CERES; ENSU, Expert-N-SUCROS; ENSP, Expert-N-SPASS; AGC, AgroC

water retention curves with field-measured suction and water content data that nonequilibrium conditions in the soil water retention dynamics can be relevant, especially in the upper soil horizons (Herbrich & Gerke, 2017). Nevertheless, incorporation of these complex processes was out of the scope of this study.

The nRMSE values of the agronomic and environmental ecosystem variables ranged between different crop models from 10 to 428%. The nRMSE values ranged from 117 to 410% for LAI, 10 to 148% for AgBio, 26 to 186% for GY, 36 to 388% for SWC, 50 to 129% for ET_a , and 95 to 428% for NetQ. This shows that the errors in simulated agronomic ecosystem variables were, in general, much larger for the in-season variables (nRMSE = 36–428%) than for the end-of-season variables. AgBio and GY (nRMSE = 10–186%). As mentioned, the in-season variables LAI and NetQ showed the largest uncertainties. The partially

large variability among crop model outputs reflects the differences in model structure and relationships, or simply the use of other model parameter values to parameterize the same process. The results agree with previous investigations for simulations of ET_a (Kimball et al., 2019), phenology (Wallach et al., 2019), and yield (Asseng et al., 2013). For the agronomic ecosystem variables, different individual crop models, mostly from the EN model family, yielded the lowest nRMSE values, but for the environmental ecosystem-related variables, no individual crop model outperformed the others in terms of nRMSE. A comparison between Richards- and non-Richards-based models showed that non-Richards-based models achieved higher nRMSE values for SWC and NetQ. The overestimation of NetQ goes along with a lower plant water availability, which indirectly limits the root water uptake and yields lower T_a . However, no clear effect on the nRMSE values

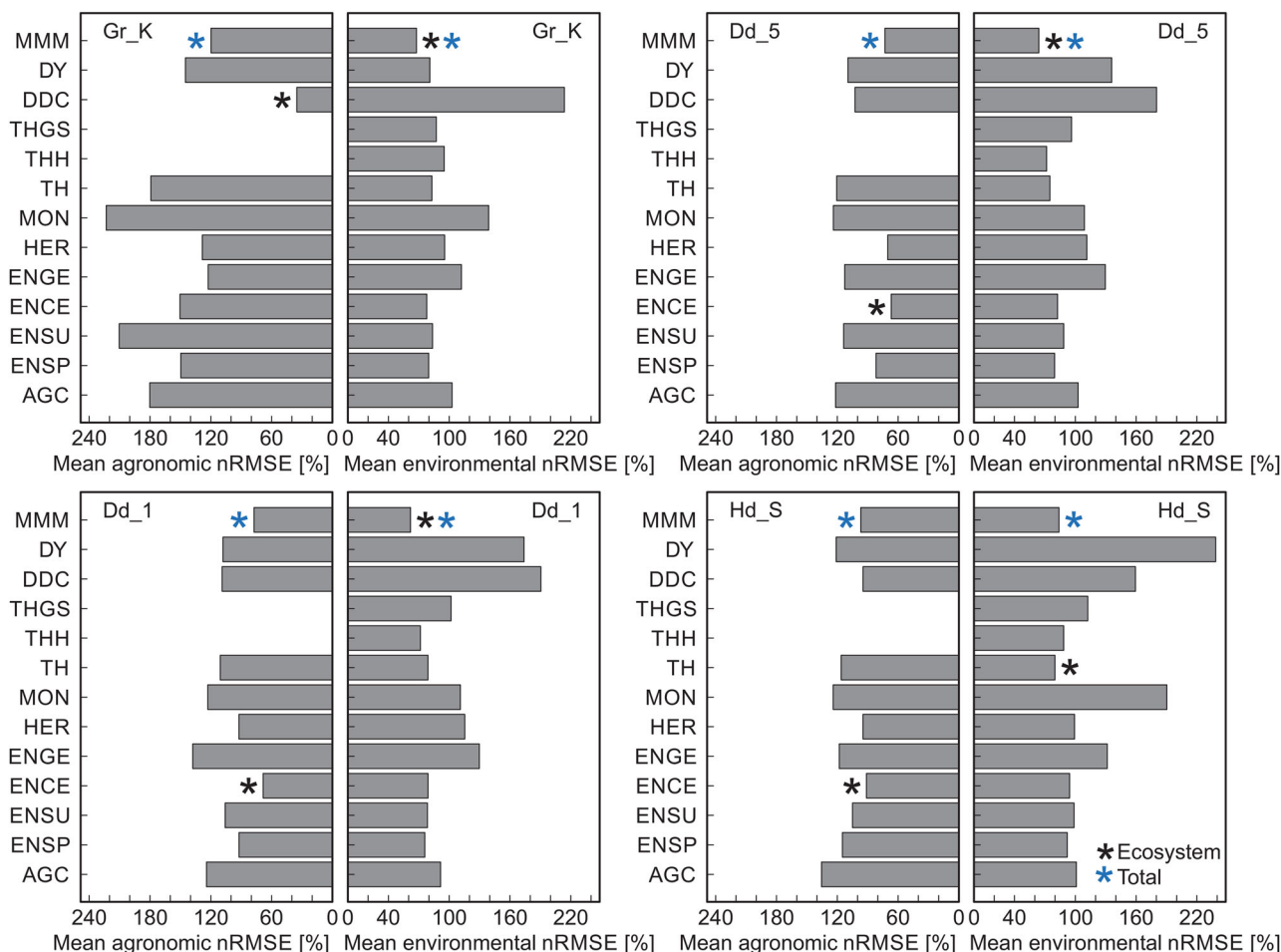


FIGURE 7 Normalized RMSE (nRMSE) between the simulated and observed averaged agroecosystem variable balances and states for all four soils (Calcaric Regosol [Gr_K], Nudiargic Luvisol [Dd_5], Calcic Luvisol [Dd_1], and Colluvic Regosol [Hd_S]), individual models, and the multi-model mean (MMM). The asterisk marks the simulation approach with the best model performance of the individual (black) and overall (blue) agroecosystem variables, respectively. Model abbreviations: DY, Daisy; DDC, DailyDayCent; THGS, THESEUS-HydroGeoSphere; THH, THESEUS-Hydrus-1D; TH, THESEUS; MON, MONICA; HER, HERMES; ENGE, Expert-N-GECROS; ENCE, Expert-N-CERES; ENSU, Expert-N-SUCROS; ENSP, Expert-N-SPASS; AGC, AgroC

could be seen for other variables (e.g., for ET_a), as here water fluxes might be compensated by the use of different root uptake functions and root distributions (Wu & Kersebaum, 2008), partitioning of ET_a , as well as water stress, which is included in the root water uptake and heat stress functions. Hence, the results emphasize the importance of how soil physics and root water uptake are represented in the different crop models.

Overall, the MMM showed the most agreement between the corresponding observation and simulation, indicating that the MMM is the best predictor. Two models, which use similar routines to simulate water fluxes (by the use of the Richards equation) and also share the same crop growth scheme (SUCROS), are AGC and ENSU. Notably, both models show the same pattern of deviations from the measured GY, even though they differ in the absolute values (Figure 4). The outputs of both models differ with

respect to LAI, AgBio, and SWC, particularly for the Calcaric Regosol (Gr_K; see Figure 6). For ET_a and NetQ, the differences between the two models are less pronounced. Given that the two models are quite similar, this could only result from using different crop parameters, which also points to the fact that the simulation results are always affected by the inherently biased choice of model parameters selected by the user (Diekkrüger, Söndgerath, Kersebaum, & McVoy, 1995; Herbst et al., 2005). Outcome from different EN models (with the exception of ENGE) showed similar nRMSE values, with respect to environmental variables, for ET_a , SWC, and NetQ. This reflects the relatively similar functionality and parameterization of the soil in the EN models. However, the use of different crop growth schemes (CERES, GECROS, SPASS, and SUCROS) in EN resulted in a larger spread of the model performance for agronomic variables (LAI, AgBio, and GY).

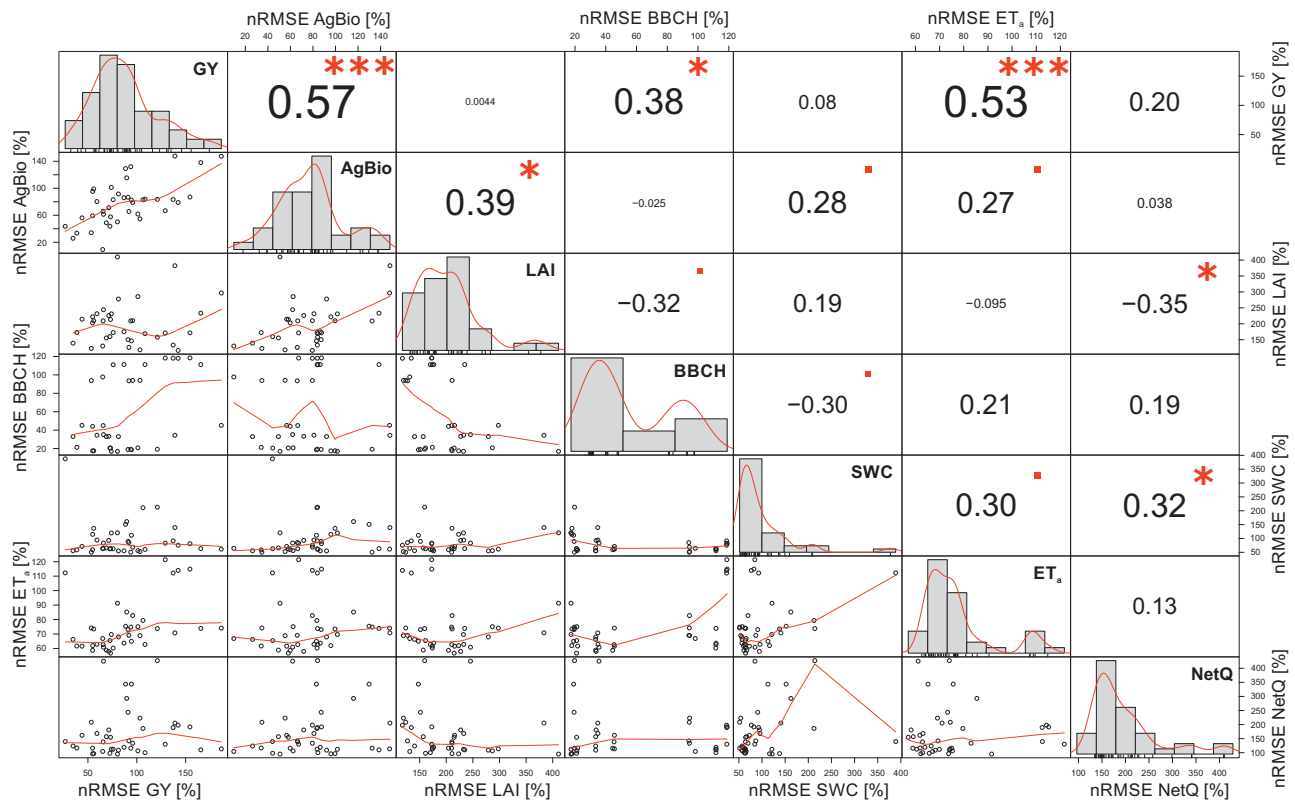


FIGURE 8 Correlation analysis between the normalized root mean square error values (nRMSE) of agronomic and environmental fluxes and states. The distribution of each variable is shown on the diagonal. The bivariate scatter plots with a fitted line are displayed on the bottom of the diagonal. On the top of the diagonal, the value of the Spearman's correlation coefficient, ρ , is also provide. Significance levels are shown by the asterisks, whereby the significance level is associated with a specific symbol: *** p value = .0001, ** p value = .001, * p value = .01. AgBio, total aboveground biomass; BBCH, phenological stages; ET_a, actual evapotranspiration; LAI, leaf area index; SWC, soil water content; NetQ, net water flux across the lysimeter bottom at 1.5-m soil depth

Figure 7 summarizes the individual model performance, pointing to the large differences among models and between the models and the MMM to predict agronomic and environmental ecosystem-related fluxes and states. For all responses, the equal-weighted mean nRMSE values of the MMM were lower than any individual average nRMSE over a specific crop model and soil. The better predictions for ET_a and NetQ by the MMM, as compared with a particular crop model, were already reported for other agronomic variables in earlier studies (Asseng et al., 2013; Martre et al., 2015; Wallach et al., 2018).

The low predictability of the observed effects by the single crop models or the MMM within this model exercise might be related to the assumed maximum rooting depth of each soil, as root development in such a specific soil landscape depends on the erosion-induced soil profile modifications such as horizon thickness, vertical sequence of horizons, and structural development of the horizon (Herbrich, Gerke, & Sommer, 2017). To account for the different potential rooting depths, the rooting depth was adjusted to the soil profile in a second step and ranged from 0.47 at the hilltop to 1.5 m at depression. Unfortunately,

the adapted maximum rooting depth, which accounts for the erosion-induced soil profile modifications, did not lead to a significant improvement of the average nRMSE of the corresponding soil profile observations. Moreover, the nRMSE averaged over all soils of the MMM increased from 81 to 92%, when accounting for soil-specific maximum rooting depth. Comparing the nRMSE values for MMM showed that an adjusted maximum rooting depth led, on average, to an improved description of the in-season variables LAI, SWC, and NetQ, but at the same time to larger mismatch of the end-of-season variables AgBio and GY. The latter might be related to a more pronounced occurrence of water and nutrient stress during the growing season.

3.2 | Relationship between the errors of agronomic and environmental fluxes and states

The analysis of the relationship between the model errors of different variables (Figure 8) revealed significant

correlations between specific in- and end-of-season variables. We found significant positive correlations (>0.5) between end-of-season variables GY and AgBio and between the in- and end-of-season variables ET_a and GY. Close correlations (>0.3 and <0.5) were identified between end-of- and in-season variables AgBio and LAI, and GY and BBCH, and between in-season variables LAI and BBCH (negative), BBCH and SWC (negative), SWC and ET_a (positive), LAI and NetQ (negative), and ET and NetQ (positive). This is in line with results from Martre et al. (2015), who showed that end-of-season variables like AgBio were highly related to GY, whereas in-season variables like SWC and LAI were solely related to the total aboveground N and the N harvest index. The correlations between GY and ET_a and between AgBio and LAI demonstrate that errors in simulating the most important end-of-season values are related to errors in simulating in-season growth processes. This suggests that in-season variables like LAI, SWC, NetQ, and especially ET_a might be of high relevance to improve predictions of crop model simulations. This is in line with Specka, Nendel, and Wieland (2019), who showed that parameters of process-based agroecosystem models were sensitive during a specific period of the growing season. Including in-season observations in the calibration helps simulating and describing more realistically in-season growth processes, which finally lead to end-of-season values of AgBio and GY (Martre et al., 2015).

4 | CONCLUSION

Our study used data from the TERENO-SOILCan lysimeter network site Dedelow to compare 12 agroecosystem models for simulating agronomic and environmental variables with limited calibration, based on phenology stages only. Simulation results and data for GY, AgBio, LAI, ET_a , drainage (NetQ), and SWC were compared with data for a period from August 2014 until October 2018.

The results of the crop model intercomparison suggest that the predictive capability of crop models is highly diverse for both agronomic and environmental variables. So far, combined analysis of model performance for different data representing agronomic and environmental variables has rarely been done in model comparison studies. The divergence between predicting LAI and ET_a suggests that the partitioning of ET_a into actual evaporation (E_a) and actual transpiration (T_a), which is of crucial importance for terrestrial water balance and for the prediction of future ecosystem feedbacks, is insufficiently described by most crop models. The weighable lysimeters observed interactions between crop yield, plant development, and water fluxes were largely controlled by erosion-affected

soil states and related soil properties. This demonstrates that soil is important when attempting to describe both the agronomic and environmental fluxes and states. However, this can be tested only where drainage data are available, such as in this weighing lysimeter study.

Simulation results clearly showed the need to calibrate crop models on a much broader range of agronomic and environmental variables, because neither the individual crop models nor the MMM were able to predict the observed representative agronomic and environmental variables (e.g., LAI, GY, ET_a , and NetQ) for the four lysimeters with truncated and colluvial soil profiles. Regardless of questions about model parameterization and model structure, the results suggest that the data on soil hydraulic properties, phenological crop stages, and boundary conditions are insufficient for calibrating crop models to represent field-scale heterogeneity in erosion-affected agricultural soil landscapes. In line with previous studies, the MMM approach was useful here, as it largely reduced the errors in predicting water fluxes.

Our analysis showed that:

1. The predictive capability of the models was highly diverse for simulating both crop development and environmental fluxes.
2. Soil does matter in agroecosystem models, and lysimeters provide such soil-related data for testing modeling of soil–vegetation–atmosphere processes.
3. Erosion- and deposition-induced changes in depth functions of soil properties are relevant in understanding biomass production, water fluxes, and soil states.
4. Differences between erosion-affected soils in crop yield, water fluxes, and states could not satisfactorily be described by individual models and the MMM when calibrated for crop BBCH only.

The results herald the need to account for spatial variability of soils within a detailed calibration procedure of the crop models when trying to improve predictions of agronomic and environmental fluxes and states. This especially holds true for erosion-affected soil landscapes of hummocky ground moraines. The varying thickness and sequence of soil horizons of the soil monoliths in the presently studied lysimeters reflect soil profiles that result from long-term effects of erosion and tillage, which is typical for intensively cultivated, post-glacial morainic soil landscapes.

ACKNOWLEDGMENTS

We acknowledge the support of TERENO and SOILCan, which were funded by the Helmholtz Association (HGF) and the Federal Ministry of Education and Research (BMBF). The TERENO-SOILCan lysimeters in Dedelow

are operated by the Leibniz Centre for Agricultural Landscape Research (ZALF) Müncheberg. Evelyn Wallor was additionally funded by the I4S Project within the BMBF BonaRes Program (031B05131). Tobias Weber was financially supported by the Collaborative Research Center 1253 CAMPOS (Project 7: Stochastic Modeling Framework), funded by the German Research Foundation (DFG). We thank Jörg Haase, Dr. Gernot Verch, Ingrid Onasch, and Gudrun Buddrus for data collection and maintenance of the experimental setup at the ZALF Research Station Dedelow.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest

ORCID

Jannis Groh  <https://orcid.org/0000-0002-1681-2850>

Efstathios Diamantopoulos  <https://orcid.org/0000-0001-7870-0291>

Michael Herbst  <https://orcid.org/0000-0003-1371-4208>

Maja Holbak  <https://orcid.org/0000-0001-5164-2327>

Bahareh Kamali  <https://orcid.org/0000-0002-8070-0175>

Kurt-Christian Kersebaum  <https://orcid.org/0000-0002-3679-8427>

Gunnar Lischeid  <https://orcid.org/0000-0003-3700-6062>


Claas Nendel  <https://orcid.org/0000-0001-7608-9097>

Eckart Priesack  <https://orcid.org/0000-0002-5088-9528>

Jörg Steidl  <https://orcid.org/0000-0002-6599-0450>

Thomas Pütz  <https://orcid.org/0000-0003-2101-448X>

Harry Vereecken  <https://orcid.org/0000-0002-8051-8517>

Tobias K.D. Weber  <https://orcid.org/0000-0002-3448-5208>

Lutz Weihermüller  <https://orcid.org/0000-0003-1991-7735>

Horst H. Gerke  <https://orcid.org/0000-0002-6232-7688>

REFERENCES

- Abrahamsen, P., & Hansen, S. (2000). Daisy: An open soil-crop-atmosphere system model. *Environmental Modelling & Software*, 15, 313–330. [https://doi.org/10.1016/S1364-8152\(00\)00003-7](https://doi.org/10.1016/S1364-8152(00)00003-7)
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration: Guidelines for computing crop water requirements*. Rome: FAO.
- Allen, R. G., Pruitt, W. O., Wright, J. L., Howell, T. A., Ventura, F., Snyder, R., ... Elliott, R. (2006). A recommendation on standardized surface resistance for hourly calculation of reference ET₀ by the FAO56 Penman-Monteith method. *Agricultural Water Management*, 81, 1–22. <https://doi.org/10.1016/j.agwat.2005.03.007>
- Aquanty. (2013). *HGS 2013: HydroGeoSphere user manual*. Waterloo, ON, Canada: Aquanty.
- Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., ... Zhu, Y. (2014). Rising temperatures reduce global wheat production. *Nature Climate Change*, 5, 143–147. <https://doi.org/10.1038/nclimate2470>
- Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., ... Wolf, J. (2013). Uncertainty in simulating wheat yields under climate change. *Nature Climate Change*, 3, 827–832. <https://doi.org/10.1038/nclimate1916>
- Asseng, S., Martre, P., Maiorano, A., Rötter, R. P., O'Leary, G. J., Fitzgerald, G. J., ... Ewert, F. (2019). Climate change impact and adaptation for wheat protein. *Global Change Biology*, 25, 155–173. <https://doi.org/10.1111/gcb.14481>
- Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J. W., ... Waha, K. (2014). How do various maize crop models vary in their responses to climate change factors? *Global Change Biology*, 20, 2301–2320. <https://doi.org/10.1111/gcb.12520>
- Battisti, R., Sentelhas, P. C., & Boote, K. J. (2017). Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. *Field Crops Research*, 200, 28–37. <https://doi.org/10.1016/j.fcr.2016.10.004>
- Battisti, R., Sentelhas, P. C., Parker, P. S., Nendel, C., Câmara, G. M. D.S., Farias, J. R. B., & Basso, C. J. (2018). Assessment of crop-management strategies to improve soybean resilience to climate change in southern Brazil. *Crop and Pasture Science*, 69, 154–162. <https://doi.org/10.1071/CP17293>
- Biernath, C., Gayler, S., Bittner, S., Klein, C., Högy, P., Fangmeier, A., & Priesack, E. (2011). Evaluating the ability of four crop models to predict different environmental impacts on spring wheat grown in open-top chambers. *European Journal of Agronomy*, 35, 71–82. <https://doi.org/10.1016/j.eja.2011.04.001>
- Brunner, P., & Simmons, C. T. (2012). HydroGeoSphere: A fully integrated, physically based hydrological model. *Groundwater*, 50, 170–176. <https://doi.org/10.1111/j.1745-6584.2011.00882.x>
- Cammarano, D., Rötter, R. P., Asseng, S., Ewert, F., Wallach, D., Martre, P., ... Wolf, J. (2016). Uncertainty of wheat water use: Simulated patterns and sensitivity to temperature and CO₂. *Field Crops Research*, 198, 80–92. <https://doi.org/10.1016/j.fcr.2016.08.015>
- Challinor, A. J., Müller, C., Asseng, S., Deva, C., Nicklin, K. J., Wallach, D., ... Koehler, A. K. (2018). Improving the use of crop models for risk assessment and climate change adaptation. *Agricultural Systems*, 159, 296–306. <https://doi.org/10.1016/j.agsy.2017.07.010>
- Chapman, M., Walker, W. S., Cook-Patton, S. C., Ellis, P. W., Farina, M., Griscom, B. W., & Baccini, A. (2020). Large climate mitigation potential from adding trees to agricultural lands. *Global Change Biology*, 26, 4357–4365. <https://doi.org/10.1111/gcb.15121>
- Coucheney, E., Buis, S., Launay, M., Constantin, J., Mary, B., García de Cortázar-Atauri, I., ... Léonard, J. (2015). Accuracy, robustness and behavior of the STICS soil-crop model for plant, water and nitrogen outputs: Evaluation over a wide range of agro-environmental conditions in France. *Environmental Modelling & Software*, 64, 177–190. <https://doi.org/10.1016/j.envsoft.2014.11.024>
- Del Grosso, S. J., Parton, W. J., Mosier, A. R., Hartman, M. D., Brenner, J., Ojima, D. S., & Schimel, D. S. (2001). Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT model. In M. J. Shaffer, L. Ma, & S. Hansen (Eds.), *Modeling carbon and nitrogen dynamics for soil management* (pp. 303–332). Boca Raton, FL: CRC Press.
- Del Grosso, S. J., Parton, W. J., Mosier, A. R., Walsh, M. K., Ojima, D. S., & Thornton, P. E. (2006). DAYCENT national-scale simulations of nitrous oxide emissions from cropped soils in the United States.

- Journal of Environmental Quality*, 35, 1451–1460. <https://doi.org/10.2134/jeq2005.0160>
- Diamantopoulos, E., Šimůnek, J., Oberdörster, C., Hammel, K., Jene, B., Schröder, T., ... Harter, T. (2017). Assessing the potential exposure of groundwater to pesticides: A model comparison. *Vadose Zone Journal*, 16. <https://doi.org/10.2136/vzj2017.04.0070>
- Diekkrüger, B., Söndgerath, D., Kersebaum, K. C., & McVoy, C. W. (1995). Validity of agroecosystem models a comparison of results of different models applied to the same data set. *Ecological Modelling*, 81, 3–29. [https://doi.org/10.1016/0304-3800\(94\)00157-D](https://doi.org/10.1016/0304-3800(94)00157-D)
- Durand, J.-L., Delusca, K., Boote, K., Lizaso, J., Manderscheid, R., Weigel, H. J., ... Zhao, Z. (2018). How accurately do maize crop models simulate the interactions of atmospheric CO₂ concentration levels with limited water supply on water use and yield? *European Journal of Agronomy*, 100, 67–75. <https://doi.org/10.1016/j.eja.2017.01.002>
- Engel, T., & Priesack, E. (1993). Expert-N: A building block system of nitrogen models as resource for advice, research, water management and policy. In H. J. P. Eijssackers & T. Hamers (Eds.), *Integrated soil and sediment research: A basis for proper protection* (pp. 503–507). Dordrecht, the Netherlands: Springer.
- Fleisher, D. H., Condori, B., Quiroz, R., Alva, A., Asseng, S., Barreda, C., ... Woli, P. (2017). A potato model intercomparison across varying climates and productivity levels. *Global Change Biology*, 23, 1258–1281. <https://doi.org/10.1111/gcb.13411>
- Frieler, K., Schauburger, B., Arneth, A., Balkovič, J., Chrysanthopoulos, J., Deryng, D., ... Levermann, A. (2017). Understanding the weather signal in national crop-yield variability. *Earth's Future*, 5, 605–616. <https://doi.org/10.1002/2016EF000525>
- Gerke, H. H., & Hierold, W. (2012). Vertical bulk density distribution in C-horizons from marley till as indicator for erosion history in a hummocky post-glacial soil landscape. *Soil and Tillage Research*, 125, 116–122. <https://doi.org/10.1016/j.still.2012.06.005>
- Gerke, H. H., Rieckh, H., & Sommer, M. (2016). Interactions between crop, water, and dissolved organic and inorganic carbon in a hummocky landscape with erosion-affected pedogenesis. *Soil and Tillage Research*, 156, 230–244. <https://doi.org/10.1016/j.still.2015.09.003>
- Groh, J., Vanderborght, J., Pütz, T., & Vereecken, H. (2016). How to control the lysimeter bottom boundary to investigate the effect of climate change on soil processes? *Vadose Zone Journal*, 15(7). <https://doi.org/10.2136/vzj2015.08.0113>
- Große, B. (1963). Über den Einfluß der Bodenerosion auf die Ertragsfähigkeit von Parabraunerden aus Löß. *Zeitschrift für Pflanzenernährung, Düngung, Bodenkunde*, 102, 212–217. <https://doi.org/10.1002/jpln.19631020304>
- Hansen, S., Abrahamsen, P., Petersen, C. T., & Styczen, M. (2012). Daisy: Model use, calibration, and validation. *Transactions of the ASABE*, 55, 1315–1334. <https://doi.org/10.13031/2013.42244>
- Hasegawa, T., Li, T., Yin, X., Zhu, Y., Boote, K., Baker, J., ... Zhu, J. (2017). Causes of variation among rice models in yield response to CO₂ examined with Free-Air CO₂ enrichment and growth chamber experiments. *Scientific Reports*, 7. <https://doi.org/10.1038/s41598-017-13582-y>
- Heinrich, I., Balanzategui, D., Bens, O., Blasch, G., Blume, T., Böttcher, F., ... Wilken, F. (2018). Interdisciplinary geo-ecological research across time scales in the Northeast German Lowland Observatory (TERENO-NE). *Vadose Zone Journal*, 17. <https://doi.org/10.2136/vzj2018.06.0116>
- Herbrich, M., & Gerke, H. H. (2017). Scales of water retention dynamics observed in eroded Luvisols from an arable postglacial soil landscape. *Vadose Zone Journal*, 16(10). <https://doi.org/10.2136/vzj2017.01.0003>
- Herbrich, M., Gerke, H. H., Bens, O., & Sommer, M. (2017). Water balance and leaching of dissolved organic and inorganic carbon of eroded Luvisols using high precision weighing lysimeters. *Soil and Tillage Research*, 165, 144–160. <https://doi.org/10.1016/j.still.2016.08.003>
- Herbrich, M., Gerke, H. H., & Sommer, M. (2017). Root development of winter wheat in erosion-affected soils depending on the position in a hummocky ground moraine soil landscape. *Journal of Plant Nutrition and Soil Science*, 181, 147–157. <https://doi.org/10.1002/jpln.201600536>
- Herbst, M., Fialkiewicz, W., Chen, T., Pütz, T., Thiéry, D., Mouvet, C., ... Vereecken, H. (2005). Intercomparison of flow and transport models applied to vertical drainage in cropped lysimeters. *Vadose Zone Journal*, 4, 240–254. <https://doi.org/10.2136/vzj2004.0070>
- Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., ... Ewert, F. (2016). Impact of spatial soil and climate input data aggregation on regional yield simulations. *PLOS ONE*, 11(4). <https://doi.org/10.1371/journal.pone.0151782>
- Hoffmann, M., Jurisch, N., Garcia Alba, J., Albiac Borraz, E., Schmidt, M., Huth, V., ... Augustin, J. (2017). Detecting small-scale spatial heterogeneity and temporal dynamics of soil organic carbon (SOC) stocks: A comparison between automatic chamber-derived C budgets and repeated soil inventories. *Biogeosciences*, 14, 1003–1019. <https://doi.org/10.5194/bg-14-1003-2017>
- Hussain, J., Khaliq, T., Ahmad, A., & Akhtar, J. (2018). Performance of four crop model for simulations of wheat phenology, leaf growth, biomass and yield across planting dates. *PLOS ONE*, 13(6). <https://doi.org/10.1371/journal.pone.0197546>
- Kersebaum, K. C. (1995). Application of a simple management model to simulate water and nitrogen dynamics. *Ecological Modelling*, 81, 145–156. [https://doi.org/10.1016/0304-3800\(94\)00167-G](https://doi.org/10.1016/0304-3800(94)00167-G)
- Kersebaum, K. C. (2007). Modelling nitrogen dynamics in soil–crop systems with HERMES. *Nutrient Cycling in Agroecosystems*, 77, 39–52. <https://doi.org/10.1007/s10705-006-9044-8>
- Kersebaum, K. C., Boote, K. J., Jorgenson, J. S., Nendel, C., Bindi, M., Frühauf, C., ... Wegehenkel, M. (2015). Analysis and classification of data sets for calibration and validation of agro-ecosystem models. *Environmental Modelling & Software*, 72, 402–417. <https://doi.org/10.1016/j.envsoft.2015.05.009>
- Kimball, B. A., Boote, K. J., Hatfield, J. L., Ahuja, L. R., Stockle, C., Archontoulis, S., ... Williams, K. (2019). Simulation of maize evapotranspiration: An inter-comparison among 29 maize models. *Agricultural and Forest Meteorology*, 271, 264–284. <https://doi.org/10.1016/j.agrformet.2019.02.037>
- Klosterhalfen, A., Herbst, M., Weihermüller, L., Graf, A., Schmidt, M., Stadler, A., ... Vereecken, H. (2017). Multi-site calibration and validation of a net ecosystem carbon exchange model for croplands. *Ecological Modelling*, 363, 137–156. <https://doi.org/10.1016/j.ecolmodel.2017.07.028>
- Kollas, C., Kersebaum, K. C., Nendel, C., Manevski, K., Müller, C., Palosuo, T., ... Wu, L. (2015). Crop rotation modelling: A European model intercomparison. *European Journal of Agronomy*, 70, 98–111. <https://doi.org/10.1016/j.eja.2015.06.007>

- Li, S., Lobb, D. A., Lindstrom, M. J., & Farenhorst, A. (2007). Tillage and water erosion on different landscapes in the northern North American Great Plains evaluated using ^{137}Cs technique and soil erosion models. *Catena*, 70, 493–505. <https://doi.org/10.1016/j.catena.2006.12.003>
- Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Adam, M., ... Bouman, B. (2015). Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions. *Global Change Biology*, 21, 1328–1341. <https://doi.org/10.1111/gcb.12758>
- Luo, Y., & Sophocleous, M. (2010). Seasonal groundwater contribution to crop-water use assessed with lysimeter observations and model simulations. *Journal of Hydrology*, 389, 325–335. <https://doi.org/10.1016/j.jhydrol.2010.06.011>
- Maharjan, G. R., Hoffmann, H., Webber, H., Srivastava, A. K., Weiermüller, L., Villa, A., ... Gaiser, T. (2019). Effects of input data aggregation on simulated crop yields in temperate and Mediterranean climates. *European Journal of Agronomy*, 103, 32–46. <https://doi.org/10.1016/j.eja.2018.11.001>
- Marin, F. R., Thorburn, P. J., Nassif, D. S. P., & Costa, L. G. (2015). Sugarcane model intercomparison: Structural differences and uncertainties under current and potential future climates. *Environmental Modelling & Software*, 72, 372–386. <https://doi.org/10.1016/j.envsoft.2015.02.019>
- Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J. W., Rötter, R. P., ... Wolf, J. (2015). Multimodel ensembles of wheat growth: Many models are better than one. *Global Change Biology*, 21, 911–925. <https://doi.org/10.1111/gcb.12768>
- Meier, U. (2001). *Entwicklungsstadien mono- und dikotiler Pflanzen*. Berlin, Braunschweig: Biologische Bundesanstalt für Land- und Forstwirtschaft.
- Miller, B. A., Koszinski, S., Hierold, W., Rogasik, H., Schröder, B., Van Oost, K., ... Sommer, M. (2016). Towards mapping soil carbon landscapes: Issues of sampling scale and transferability. *Soil and Tillage Research*, 156, 194–208. <https://doi.org/10.1016/j.still.2015.07.004>
- Nendel, C., Berg, M., Kersebaum, K. C., Mirschel, W., Specka, X., Wegehenkel, M., ... Wieland, R. (2011). The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecological Modelling*, 222, 1614–1625. <https://doi.org/10.1016/j.ecolmodel.2011.02.018>
- Nendel, C., Kersebaum, K. C., Mirschel, W., & Wenkel, K. O. (2014). Testing farm management options as climate change adaptation strategies using the MONICA model. *European Journal of Agronomy*, 52, 47–56. <https://doi.org/10.1016/j.eja.2012.09.005>
- O'Leary, G. J., Christy, B., Nuttall, J., Huth, N., Cammarano, D., Stöckle, C., ... Asseng, S. (2015). Response of wheat growth, grain yield and water use to elevated CO_2 under a Free-Air CO_2 Enrichment (FACE) experiment and modelling in a semi-arid environment. *Global Change Biology*, 21, 2670–2686. <https://doi.org/10.1111/gcb.12830>
- Palosuo, T., Kersebaum, K. C., Angulo, C., Hlavinka, P., Moriondo, M., Olesen, J. E., ... Reimund, R. (2011). Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. *European Journal of Agronomy*, 35, 103–114. <https://doi.org/10.1016/j.eja.2011.05.001>
- Papiernik, S. K., Lindstrom, M. J., Schumacher, J. A., Farenhorst, A., Stephens, K. D., Schumacher, T. E., & Lobb, D. A. (2005). Variation in soil properties and crop yield across an eroded prairie landscape. *Journal of Soil and Water Conservation*, 60, 388–395.
- Parton, W. J., Holland, E. A., Del Grosso, S. J., Hartman, M. D., Martin, R. E., Mosier, A. R., ... Schimel, D. S. (2001). Generalized model for NO_x and N_2O emissions from soils. *Journal of Geophysical Research: Atmospheres*, 106, 17403–17419. <https://doi.org/10.1029/2001JD900101>
- Pennock, D., Bedard-Haughn, A., Kiss, J., & van der Kamp, G. (2014). Application of hydrogeology to predictive mapping of wetland soils in the Canadian Prairie Pothole region. *Geoderma*, 235–236, 199–211. <https://doi.org/10.1016/j.geoderma.2014.07.008>
- Pennock, D., Yates, T., Bedard-Haughn, A., Phipps, K., Farrell, R., & McDougal, R. (2010). Landscape controls on N_2O and CH_4 emissions from freshwater mineral soil wetlands of the Canadian Prairie Pothole region. *Geoderma*, 155, 308–319. <https://doi.org/10.1016/j.geoderma.2009.12.015>
- Peters, A., Groh, J., Schrader, F., Durner, W., Vereecken, H., & Pütz, T. (2017). Towards an unbiased filter routine to determine precipitation and evapotranspiration from high precision lysimeter measurements. *Journal of Hydrology*, 549, 731–740. <https://doi.org/10.1016/j.jhydrol.2017.04.015>
- Priesack, E., Gayler, S., & Hartmann, H. P. (2006). The impact of crop growth sub-model choice on simulated water and nitrogen balances. *Nutrient Cycling in Agroecosystems*, 75, 1–13. <https://doi.org/10.1007/s10705-006-9006-1>
- Pütz, T., Kiese, R., Wollschläger, U., Groh, J., Rupp, H., Zacharias, S., ... Vereecken, H. (2016). TERENO-SOILCan: A lysimeter-network in Germany observing soil processes and plant diversity influenced by climate change. *Environmental Earth Sciences*, 75. <https://doi.org/10.1007/s12665-016-6031-5>
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Raatz, L., Bacchi, N., Pirhofer Walzl, K., Glemnitz, M., Müller, M. E. H., Joshi, J., ... Scherber, C. (2019). How much do we really lose?: Yield losses in the proximity of natural landscape elements in agricultural landscapes. *Ecology and Evolution*, 9, 7838–7848. <https://doi.org/10.1002/ece3.5370>
- Rezaei, E. E., Siebert, S., Manderscheid, R., Müller, J., Mahrookashani, A., Ehrenpfordt, B., ... Ewert, F. (2018). Quantifying the response of wheat yields to heat stress: The role of the experimental setup. *Field Crops Research*, 217, 93–103. <https://doi.org/10.1016/j.fcr.2017.12.015>
- Rieckh, H., Gerke, H. H., Siemens, J., & Sommer, M. (2014). Water and dissolved carbon fluxes in an eroding soil landscape depending on terrain position. *Vadose Zone Journal*, 13(7). <https://doi.org/10.2136/vzj2013.10.0173>
- Rieckh, H., Gerke, H. H., & Sommer, M. (2012). Hydraulic properties of characteristic horizons depending on relief position and structure in a hummocky glacial soil landscape. *Soil and Tillage Research*, 125, 123–131. <https://doi.org/10.1016/j.still.2012.07.004>
- Rötter, R. P., Palosuo, T., Kersebaum, K. C., Angulo, C., Bindi, M., Ewert, F., ... Trnka, M. (2012). Simulation of spring barley yield in different climatic zones of Northern and Central Europe: A comparison of nine crop models. *Field Crops Research*, 133, 23–36. <https://doi.org/10.1016/j.fcr.2012.03.016>
- Salo, T. J., Palosuo, T., Kersebaum, K. C., Nendel, C., Angulo, C., Ewert, F., ... Rötter, R. P. (2016). Comparing the performance of 11 crop simulation models in predicting yield response to nitrogen fertilization. *The Journal of Agricultural Science*, 154, 1218–1240. <https://doi.org/10.1017/S0021859615001124>

- Schauberger, B., Archontoulis, S., Arneith, A., Balkovic, J., Ciaia, P., Deryng, D., ... Frieler, K. (2017). Consistent negative response of US crops to high temperatures in observations and crop models. *Nature Communications*, 8. <https://doi.org/10.1038/ncomms13931>
- Schils, R., Olesen, J. E., Kersebaum, K.-C., Rijk, B., Oberforster, M., Kalyada, V., ... van Ittersum, M. K. (2018). Cereal yield gaps across Europe. *European Journal of Agronomy*, 101, 109–120. <https://doi.org/10.1016/j.eja.2018.09.003>
- Seidel, S. J., Palosuo, T., Thorburn, P., & Wallach, D. (2018). Towards improved calibration of crop models: Where are we now and where should we go? *European Journal of Agronomy*, 94, 25–35. <https://doi.org/10.1016/j.eja.2018.01.006>
- Šimůnek, J., van Genuchten, M. T., & Šejna, M. (2016). Recent developments and applications of the HYDRUS computer software packages. *Vadose Zone Journal*, 15(7). <https://doi.org/10.2136/vzj2016.04.0033>
- Sommer, M., Fiedler, S., Glatzel, S., & Kleber, M. (2004). First estimates of regional (Allgäu, Germany) and global CH₄ fluxes from wet colluvial margins of closed depressions in glacial drift areas. *Agriculture, Ecosystems & Environment*, 103, 251–257. <https://doi.org/10.1016/j.agee.2003.09.019>
- Sommer, M., Gerke, H. H., & Deumlich, D. (2008). Modelling soil landscape genesis: A “time split” approach for hummocky agricultural landscapes. *Geoderma*, 145, 480–493. <https://doi.org/10.1016/j.geoderma.2008.01.012>
- Specka, X., Nendel, C., & Wieland, R. (2019). Temporal sensitivity analysis of the MONICA Model: Application of two global approaches to analyze the dynamics of parameter sensitivity. *Agriculture*, 9(2). <https://doi.org/10.3390/agriculture9020037>
- Verch, G., Kächele, H., Höltl, K., Richter, C., & Fuchs, C. (2009). Comparing the profitability of tillage methods in northeast Germany: A field trial from 2002 to 2005. *Soil and Tillage Research*, 104, 16–21. <https://doi.org/10.1016/j.still.2008.12.012>
- Wallach, D. (2006). Evaluating crop models. In D. Wallach, D. Makowski, & J. W. Jones (Eds.), *Working with dynamic crop models: Evaluation, analysis, parameterization, and applications* (1st ed., pp. 11–54). Amsterdam: Elsevier.
- Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P. J., ... Zhang, Z. (2018). Multimodel ensembles improve predictions of crop–environment–management interactions. *Global Change Biology*, 24, 5072–5083. <https://doi.org/10.1111/gcb.14411>
- Wallach, D., Palosuo, T., Thorburn, P., Seidel, S. J., Gourdain, E., Asseng, S., ... Zhu, Y. (2019). How well do crop models predict phenology, with emphasis on the effect of calibration? *bioRxiv*. <https://doi.org/10.1101/708578>
- Waller, E., Kersebaum, K.-C., Ventrella, D., Bindi, M., Cammarano, D., Coucheney, E., ... Trombi, G. (2018). The response of process-based agro-ecosystem models to within-field variability in site conditions. *Field Crops Research*, 228, 1–19. <https://doi.org/10.1016/j.fcr.2018.08.021>
- Wang, E., & Engel, T. (1998). Simulation of phenological development of wheat crops. *Agricultural Systems*, 58, 1–24. [https://doi.org/10.1016/S0308-521X\(98\)00028-6](https://doi.org/10.1016/S0308-521X(98)00028-6)
- Wegehenkel, M., Luzi, K., Sowa, D., Barkusky, D., & Mirschel, W. (2019). Simulation of long-term soil hydrological conditions at three agricultural experimental field plots compared with measurements. *Water*, 11(5). <https://doi.org/10.3390/w11050989>
- Wöhling, T., Gayler, S., Priesack, E., Ingwersen, J., Wizemann, H.-D., Högy, P., ... Streck, T. (2013). Multiresponse, multiobjective calibration as a diagnostic tool to compare accuracy and structural limitations of five coupled soil-plant models and CLM3.5. *Water Resources Research*, 49, 8200–8221. <https://doi.org/10.1002/2013WR014536>
- Wöhling, T., Schöninger, A., Gayler, S., & Nowak, W. (2015). Bayesian model averaging to explore the worth of data for soil-plant model selection and prediction. *Water Resources Research*, 51, 2825–2846. <https://doi.org/10.1002/2014WR016292>
- Wu, L., & Kersebaum, K. C. (2008). Modeling water and nitrogen interaction responses and their consequences in crop models. In L. R. Ahuja, V. R. Reddy, S. A. Saseendran, & Q. Yu (Eds.), *Response of crops to limited water: Understanding and modeling water stress effects on plant growth processes* (pp. 215–249). Madison, WI: ASA, CSSA, and SSSA.
- Yeluripati, J. B., van Oijen, M., Wattenbach, M., Neftel, A., Ammann, A., Parton, W. J., & Smith, P. (2009). Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. *Soil Biology and Biochemistry*, 41, 2579–2583. <https://doi.org/10.1016/j.soilbio.2009.08.021>
- Yin, X., Kersebaum, K. C., Kollas, C., Baby, S., Beaudoin, N., Manevski, K., ... Olesen, J. (2017). Multi-model uncertainty analysis in predicting grain N for crop rotations in Europe. *European Journal of Agronomy*, 84, 152–165. <https://doi.org/10.1016/j.eja.2016.12.009>
- Yin, X., Kersebaum, K. C., Kollas, C., Manevski, K., Baby, S., Beaudoin, N., ... Olesen, J. (2017). Performance of process-based models for simulation of grain N in crop rotations across Europe. *Agricultural Systems*, 154, 63–77. <https://doi.org/10.1016/j.agsy.2017.03.005>
- Zambrano-Bigiarini, M. (2017). *hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series*. R package version 0.3-10. Mauricio Zambrano-Bigiarini. Retrieved from <http://hzambran.github.io/hydroGOF/>
- Żydelis, R., Weihermüller, L., Herbst, M., Klosterhalfen, A., & Lazauskas, S. (2018). A model study on the effect of water and cold stress on maize development under nemoral climate. *Agricultural and Forest Meteorology*, 263, 169–179. <https://doi.org/10.1016/j.agrformet.2018.08.011>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Groh J, Diamantopoulos E, Duan X, et al. Crop growth and soil water fluxes at erosion-affected arable sites: Using weighing lysimeter data for model intercomparison. *Vadose Zone J.* 2020;19:e20058. <https://doi.org/10.1002/vzj2.20058>