



Generic optimization approach of soil hydraulic parameters for site-specific model applications

Jonas Trenz¹ · Emir Memić¹ · William D. Batchelor² · Simone Graeff-Hönninger¹

Accepted: 16 October 2023 / Published online: 11 November 2023
© The Author(s) 2023

Abstract

Site-specific crop management is based on the postulate of varying soil and crop requirements in a field. Therefore, a field is separated into homogenous management zones, using available data to adapt management practices environment to maximize productivity and profitability while reducing environmental impacts. Due to advancing sensor technologies, crop growth and yield data on more minor scales are common, but soil data often needs to be more appropriate. Crop growth models have shown promise as a decision support tool for site-specific farming. The Decision Support System for Agrotechnology Transfer (DSSAT) is a widely used point-based model. To overcome the problem of inappropriate soil input data problem, this study introduces an external plug-in program called Soil Profile Optimizer (SPO), which uses the current DSSAT v4.8 to calibrate soil profile parameters on a site-specific level. Developed as an inverse modelling approach, the SPO can calibrate selected soil profile parameters by targeting available in-season plant data. Root Mean Square Error (RMSE) and normalized RMSE as error minimization criteria are used. The SPO was tested and evaluated by comparing different simulation scenarios in a case study of a 3-yr field trial with maize. The scenario with optimized soil profiles, conducted with the SPO, resulted in an R^2 of 0.76 between simulated and observed yield and led to significant improvements compared to the scenario conducted with field scale soil profile information (R^2 0.03). The SPO showed promise in using spatial plant measurements to estimate management zone scale soil parameters required for the DSSAT model.

Keywords Precision agriculture · Crop models · Site-specific management · Inverse modelling · Soil profile parameters optimization · Hydraulic parameters

Introduction

The approach of Precision Agriculture (PA) and its potential to increase yield in agricultural production or reduce environmental impacts and input costs by variable-rate application of production inputs was introduced in the mid-1980s. Advancing technological

✉ Jonas Trenz
j.trenz@uni-hohenheim.de

¹ Agronomy, University of Hohenheim, Stuttgart, Germany

² Biosystems Engineering Department, Auburn University, Auburn, USA

progress created the conditions for sustainable and environmentally sound agricultural practices (Basso et al. 2017; Gebbers & Adamchuk, 2010; Zhang et al. 2002). The concept of PA in crop production is based on identifying spatial yield variability within a field by using average field yield as a reference to delineate high and low-yielding areas. Spatial variability is considered the result of complex interactions among site-specific characteristics like rooting depth, soil heat balance, water and oxygen balance, nutrient supply, weather, pests, and management during the growing season (Maestrini & Basso, 2018; Thorp et al. 2008). Temporal variability is the variation in yield observed for a specific field over multiple years (Basso et al. 2017). Considering both approaches enables the explanation of yield variability by interpreting temporal variability through weather-related variabilities and spatial variability through variable soil-related properties (Maestrini & Basso, 2018). Site-specific analysis can lead to spatial management decisions such as variable N application and seeding rate (Memic et al. 2019). However, to implement a systematic approach for site-specific management, current crop conditions must be provided by process-based modelling or in-season observations from remote sensing.

Process-based crop growth models are designed to quantify crop yield and yield-limiting factors while fully capturing interactions between crops and the environment (Boote et al. 1997). They are used to predict spatial variability of yield and to support the decision-making process of optimal timing of management practices (Batchelor et al. 2002; Braga & Jones, 2004). The Decision Support System for Agrotechnology Transfer (DSSAT) is one of the most widely used process-oriented crop modelling software solutions considering different management and environmental conditions within homogeneous land units for more than 40 crops (Hoogenboom et al. 2019). The DSSAT was designed based on a modular structure approach with the Cropping System Model (CSM) as its key component. This programming code includes primary agronomic components such as soil, weather, and crop management practices for simulating crop growth (Boote, 2019). Since DSSAT is a point-based model, it can be applied to site-specific areas within a field if appropriate inputs are available at that spatial scale. This can be done by dividing heterogeneous fields into smaller and relatively homogenous site-specific units that can be treated uniformly (Batchelor et al. 2002; Paz et al. 1999).

Spatial yield variability in a field can be linked to a certain extent to the variability of soil properties (Thorp et al. 2008). Soil sampling is expensive and labor intensive, and from a site-specific perspective, it is impractical for many sites within a field. Pedo transfer functions (PTFs) are commonly used to reduce the effort in soil sampling by deriving the target model input parameters from minimum input data (Bouma, 1992). PTFs use easily identifiable soil properties, like texture or porosity, to determine model inputs such as hydraulic, thermal, biochemical, and solute transport parameters. The correlation between properties and parameters is conventionally derived with regression algorithms/techniques (statistical/neural regression), calibrated using a specific region database (van Looy et al. 2017; Aitkenhead et al. 2016). However, due to their empirical nature, PTFs are generally accurate for a specific region but may be inaccurate for specific sites (Patil & Singh, 2016; Wösten et al. 2001).

A key parameter in crop production and a driver of yield variability is the plant available water capacity (PAWC) of the soil (Hoffmann et al. 2016; Maestrini & Basso, 2018; Wu et al. 2019). Physical soil properties regulate water retention, rate of water flow, the fate of nutrients, chemicals, and pollutants in soil, and determine the accessibility of water for plant uptake, crop growth, and environmental quality and, consequently, are crucial inputs for crop models (Indoria et al. 2020; Ritchie, 1998). Hoffmann et al. (2016) showed that 81% of spatial variability of yield (silage maize) could be explained by four variables: precipitation, PAWC of the

soil profile, soil profile depth, and PAWC of the topsoil. Moreover, PAWC was the dominant variable, explaining 58% of the yield variability (Hoffmann et al. 2016).

The tipping bucket approach is the general soil water balance approach used in DSSAT-CSM. The approach assumes one-dimensional water flow (Ritchie, 1998) involves integrating surface water flow, the variability of infiltration amount, and tile drainage characteristics to compute daily water available in each soil layer (Gijssman et al. 2002). The PTF developed by Saxton et al. (1986) using multilinear regression of the data collected by Rawls et al. (1982) can compute the three hydraulic parameters (Estimands): volumetric water content at drained upper limit (DUL), volumetric water content at lower limit (LL), volumetric water content at saturation (SAT) with the minimum input of the two variables of soil texture: silt and clay (Gijssman et al. 2002; Saxton et al. 1986).

Complex spatial patterns of physical soil properties combined with rainfall patterns can lead to highly variable plant water availability and rooting characteristics that affect dynamic interactions in the simulation of growth processes (Batchelor et al. 2002; Paz et al. 1999). Missing or inaccurate soil input data might lead to errors in modelling and final yield simulation (Hoffmann et al. 2016). Particularly in site-specific modelling applications, incorrect estimation of soil hydraulic parameters complicates the explanation of yield gaps and the estimation of best spatial management practices (Braga & Jones, 2004). The conclusion from the literature is that yield variability in a field among multiple crops can be mainly attributed to the spatial variability of PAWC (Batchelor & Paz, 1997; Braga & Jones, 2004; Hoffmann et al. 2016; Wu et al. 2019).

However, the precision of PTFs varies, and a generic approach to optimize soil profile parameters is recommended (van Looy et al. 2017). Several studies in crop modelling reported the impact and the advantage of using optimization to estimate soil profile properties (Batchelor et al. 2002, 2004a; Batchelor & Paz, 1997; Braga & Jones, 2004; Thorp et al. 2008; Wu et al. 2019). However, it is essential to note that the study referenced (Batchelor et al. 2002) was conducted in the early 2000s using DSSAT version 3.7. Since then, no generic optimization approach for soil profile parameters in DSSAT-CSM has been undertaken.

This study aimed to develop a generic optimization approach to estimate standard soil parameters required by DSSAT v 4.8 and future versions. The approach was based on *inverse modelling*, which optimizes the objective function (loss function) based on parameter search techniques using variables more easily measured. A software plug-in called Soil Profile Optimizer (SPO) was developed to optimize parameters required by the DSSAT-CSM model based on easily measured target variables such as yield, tops weight, leaf area index (LAI), etc. The SPO targets selected soil parameters by minimizing the difference between simulated and observed output variable/s (e.g., yield, LAI, tops weight, etc.) based on the normalized root mean square error (nRMSE).

The detailed objectives of this paper were (a) to develop a systematic approach to optimize soil profile inputs for DSSAT v 4.8 and incorporate it into a software plug-in SPO, (b) to test the application of the SPO on a 3-yr field trial dataset of maize (*Zea mays* L.) by comparing different simulation approaches, (c) to evaluate the influence of different soil profile parameters in this field trial.

Materials and methods

Experimental data

The data set used in this publication was derived from a 3-yr field trial (2006–2008) that tested different nitrogen management strategies in corn (*Zea mays* L., cultivar ‘Companero’) by evaluating the means of corn grain yield and marginal net return. The study was conducted on the Riech field at the Research Station Ihinger Hof of the University of Hohenheim, Southwestern Germany. Ihinger Hof (48.74°N, 8.93°E) is located 475 m above sea level, and the climatic conditions at the research station are characterized by a mean annual precipitation of 694 mm, a mean temperature of 8.4 °C, and a mean daily solar radiation of 10.9 MJ m⁻² (1976–2005) (Fig. 1).

The soil’s mean pH and organic matter content were 7.2, respectively, 2.6%. The nitrogen application strategy used in this work was based on a uniform application of 160 kg N ha⁻¹ as KAS (26% N). The input data used for the model input files were based on the publications of Link et al. (2013) and Memic et al. (2019). The site was ploughed (0.25 m) in autumn after the harvest of the previous crop, corn. Seedbed preparation was done shortly before sowing in April using a harrow combined with a land packer. Corn was planted at the end of April and beginning of May (Day of sowing: 26/04/2006, 26/04/2007, 07/05/2008) with a seeding rate of 9.5 kernels m⁻² and a row distance of 0.75 m. Before sowing and after harvest, soil mineral nitrogen (N_{min}) analysis was conducted at three depths (0–30 cm, 30–60 cm, 60–90 cm). The uniform nitrogen application was performed with a pneumatic fertilizer spreader (Rauch Aero 1112, Sinzheim, Germany). Pesticides were broadcast at relevant stages based on common agricultural practices. Maize was harvested by the end/mid of October each year. Grain yield was measured with a yield monitor implemented on a combine harvester.

The Riech field is 10 ha and was divided into 80 site-specific management grids (0.125 ha). Further details on the field trial are given in Link et al. (2013). The soil of the experimental site was characterized as a heavy calcareous brown earth soil with high clay content. An overview of the texture categories of the 20 experimental grids is shown in Fig. 2, the soil texture triangle created after the USDA (2019). This data set was selected because it can demonstrate different characteristics for the features of the SPO.

The Riech field study reflects a monocropping system of maize over three years. For each year, the maize cultivar and the grid pattern did not change, and the crop

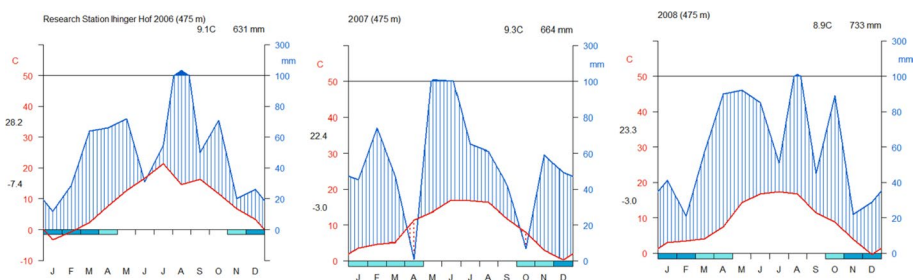


Fig. 1 Ihinger Hof Research Station average temperature and precipitation, years 2006–2008 (after Hartmann et al. (2018))

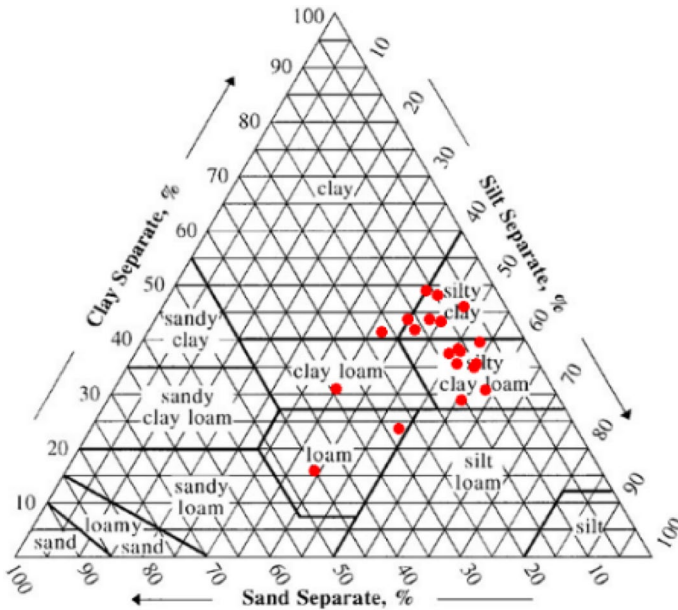


Fig. 2 Soil texture triangle determined in 20 grids (red dots) of the experimental site Riech after the soil classification of the USDA (2019) (Color figure online)

management was uniform. The experimental data of the Riech field was available on a grid-specific level for 20 grids for the years 2006, 2007, and 2008.

Soil water balance and determination of hydraulic parameters in DSSAT-CSM

Within DSSAT-CSM, daily soil water content is simulated to compute crop water stress, enabling yield prediction, crop management decision-making, risk analysis, strategic planning, and policy analysis (Boote, 2019). The soil water balance is estimated daily as a function of precipitation, irrigation, transpiration, soil evaporation, runoff, and drainage from the soil profile (Ritchie, 1998). The soil profile is divided into several computational layers, up to a maximum of 20. For simplification, a one-dimensional water flow is assumed using the tipping bucket approach developed by Ritchie (1985). The tipping bucket approach assumes that each soil layer can be filled with water up to the point of saturation (SAT) (Ritchie, 1985), defined as the maximum water content able to be held by the soil based on porosity (Godwin et al. 1984). If the volumetric water content reaches the point of saturation, the excess water drains into the next lower layer. The five hydraulic parameters: soil bulk density (BD), total porosity (TO), DUL, LL and SAT can be measured or estimated using a PTF, such as the functions developed by Saxton et al. (1986) and adapted by Gijsman et al. (2002).

The plant available water per soil layer is determined as the difference between the drained upper limit (DUL) and the lower limit (LL) in each soil layer. The DUL is the amount of water soil holds against gravity, and LL is the extent to which roots can extract water from a particular soil type (Godwin et al. 1984). The LL refers to the wilting point and thus to water potentials of -15 bar. DUL corresponds to the field

water capacity concept and water potentials in -0.1 to -0.33 bar (Ritchie, 1985). The calculation of SAT considers that some pores include entrapped air at the saturation point. The percentage of entrapped air refers to the soil type and is set for 2–3% for clay soils and up to 7% for sandy soils (Dalglish & Foale, 1998). The bulk density (BD) expresses the relationship between the bulk density of organic matter and the bulk density of mineral matter per soil volume unit. It can be calculated by the function developed by Adams (1973). Gijsman et al. (2007) classified soil data input requirements for a daily time-step into four groups: general data, apply to entire profile, first tier, and second tier. Parameters classified as available data and applied to the whole profile are used to calculate data of the first tier (e.g., LL, DUL, and SAT) and the second tier.

Analog to the definition of van Looy et al. (2017) and Vereecken et al. (2016), the two groups of general data that apply to the entire profile correspond to the definition of “Predictor,” while the first and second tier corresponds to the purpose of “Estimands”. Regarding these definitions, the hydraulic parameters used in this study are considered Estimands. To derive the Estimands, a PTF is necessary. The PTF used in DSSAT-CSM is designed to be an efficient approach to enable the modelling of the whole moisture range of the potential soil water characteristic based on specific Predictor data. Therefore, the derived equation set estimates generalized soil-water characteristics from soil texture based on the statistical correlation between soil texture and hydraulic conductivity (Saxton et al. 1986). The general approach described in Eq. 1 is based on the study of Rawls et al. (1982).

$$\Theta_p = a + b(\%sand) + c(\%silt) + d(\%clay) + e(\%organicmatter) + f\left(bulkdensity, \frac{mg}{m^3}\right) \quad (1)$$

The particle size distribution and the soil texture are defined by the USDA system (sand = 2.0–0.05 mm, silt = 0.05–0.002 mm and clay < 0.002 mm) and grouped into 12 generic soil types (Gijsman et al. 2002). The regression coefficients a, b, c, d, e, and f are determined using a stepwise multiple linear regression (Rawls et al. 1982).

The approach of Campbell (1974) in Eq. 2 describes the relationship between soil-water potential and water content.

$$\Psi = A\Theta^B \quad (2)$$

where Ψ is the soil water potential (kPa), Θ is the soil water content (m^3/m^3), coefficients A and B are fitted values. Rawls et al. (1982) demonstrated the significance of the soil texture factor clay. The Saxton approach unifies Eq. 1 and Eq. 2 to predict the soil water retention curve from particle size distribution (Saxton et al. 1986). The Saxton approach fits the data to the water retention curve by intercepting the curve into three parts and the corresponding three equations: (i) saturation to air entry constant, constant (ii) from air entry to 10 kPa linear, and (iii) from 10 to 1500 kPa curvilinear (Eqs. 3 and 4).

$$A = \exp[-4.396 - 0.0715(\%clay) - 4.880 \times 10^{-4}(\%sand)^2 - 4.285 \times 10^{-5}(\%sand)^2(\%clay)]100 \quad (3)$$

$$B = -3.140 - 0.00222(\%clay)^2 - 3.484 \times 10^{-5}(\%sand)^2(\%clay) \quad (4)$$

Saxton et al. (1986) derived Eqs. 3 and 4 out of stepwise multiple nonlinear regression. The derived coefficients A and B are unified in Eq. 2 and represent the curvilinear part from 10 to 1500 kPa of the water retention curve. The curvilinear part covers a wide range of the water retention curve and enables the determination of the PAWC. Critical

hydraulic parameters like DUL, LL and SAT are determined with equations unifying different approaches to meet the unique requirements of the DSSAT soil model.

The soil profile optimizer (SPO) - a generic algorithm for soil profile calibration

Inverse modelling is a mathematical approach used to estimate unknown parameters of a system based on observed data (Abbaspour et al. 2000). Models, or equations that describe the system's behavior, are optimized to minimize error between simulated and observed data. Inverse modeling aims to determine the best set of parameters that can reproduce the observed data. The process of inverse modelling can be separated into several steps. First, the model that describes the system's behavior is defined. Typically, the model includes parameters that need to be estimated or optimized. Next, selected parameters in the model are optimized by adjusting the values of the parameters chosen to minimize the error between simulated and observed data. This optimization technique uses a numerical optimization algorithm (Marquardt, 1963) to search for optimum parameter values. Determining soil hydraulic parameters by inverse modelling is widely used when model inputs are unknown and expensive to collect (Abbaspour et al. 2000; Kamali & Zand-Parsa, 2016; Salahou et al. 2022).

The Soil Profile Optimizer (SPO) was developed as an external software plug-in for the current DSSAT-CSM v. 4.8 based on a generic algorithm written in Python with an intuitive interface. The SPO uses inverse modelling to minimize the value of an objective function (i.e., the error between simulated and observed values). In this study, the SPO was used to optimize seven soil profile parameters required in DSSAT (Table 1). Concerning the subdivision into Predictors and Estimands, optimizing parameters of the entire soil profile (Predictors) and layer-based parameters (Estimands) is possible. Thorp et al. (2008) developed a decision support system prototype called Apollo to analyze precision farming data sets using DSSAT version 3.5. Among other things, Apollo enabled the calibration of 10 soil-related parameters to simulate historical yield (Batchelor et al. (2004) Thorp et al. 2008). Apollo is no longer supported or available. Table 1 compares the available soil model parameters in Apollo and the SPO. In Apollo, not all ten soil-related profile parameters were available for optimization. Based on the classification of the soil parameters as Predictors, the soil calibration conducted with Apollo relied on the entire soil profile parameters. The SPO enables optimizing soil profile parameters for the whole soil profile as a layer-based (e.g., SLLL throughout all defined soil layers).

The approach in the SPO focused on the soil water balance throughout all defined soil layers. Besides the general importance of the PAWC in crop production, the SPO optimization approach aims to reduce the uncertainty present in the PTF approach while deriving specific water-holding capacity properties indirectly based on measurable aspects of plant growth. Therefore, the soil hydraulic parameters 6) SLLL and 7) SDUL in Table 1 were the focus of this study. The Saxton approach was developed using 5320 soil samples valid for the USA and calculated the hydraulic parameters LL and DUL using the soil properties of clay and sand (Saxton et al. 1986). Several studies reported the challenge of accurately applying PTFs outside their development areas (Patil & Singh, 2016; Wösten et al. 2001). Furthermore, Wösten et al. (2001) investigated the accuracy and reliability of PTFs and reported an RMSE for the volumetric water contents ranging from 0.02 to 0.11 m³ m⁻³.

Table 1 Comparison of soil model parameters, including the soil data requirements for a daily time step crop simulation that can be optimized with Apollo (Thorp et al. 2008) and the SPO (after Gijssman et al. 2007)

	Apollo	SPO
Classification	Parameter name	Parameter name
Predictors	1. S LRO SCS curve number [unitless]	1. SLRO Runoff curve number [unitless]
	2. SLDR Drainage rate [Fraction/day]	2. SLDR Drainage rate [Fraction/ day]
	3. TDFD Effective Tile Drainage Rate [1/day]	–
	4. Saturated hydraulic conductivity of deep impermeable layer [cm/day]	3. SSKS Saturated hydraulic conductivity [cm h ⁻¹]
	5. Hard pan factor [0–1 scale]*	–
	6. Depth to hard pan [cm]*	–
	7. Root distribution reduction factor*	–
	8. Nitrogen mineralization factor [0–1 scale]*	–
	9. Soil fertility factor [0–1]*	4. SLPF Fertilization factor [0–1 scale]
	10. Adjust available soil water % *	–
Estimands	–	5. SRGF Root growth factor [0–1 scale]
	–	6. SLLL Lower Limit of plant extractable water [cm ³ cm ⁻³]
	–	7. SDUL Drained upper limit [cm ³ cm ⁻³]

Mechanism of the SPO

The approach in the SPO was built on the error minimization between observed and simulated target variables by calibrating selected soil model parameters. The SPO can use time-series in-season observations of specific crop model output target variables such as GWAD (Grain weight), CWAD (tops weight), LWAD (leaf weight), SWAD (stem weight), and LAI (leaf area index), observable in the field to indirectly estimate soil related parameters that might be responsible for the given variability in observed biomass and grain yield. Depending on the number of target variables and in-season observations included in the optimization procedure, the algorithm uses the RMSE or nRMSE as the error minimization criteria. In the case of one target variable and one word per season, RMSE is chosen. If there are multiple target variables with multiple in-season observations, the algorithm relies on the nRMSE. The selection of the nRMSE as the primary error minimization method enables the soil profile parameter optimization using target variables with different unit scales, e.g., LAI (leaf area m^2 per ground area m^2), GWAD (kg ha^{-1}), total above-ground biomass (CWAD) (kg ha^{-1}) (Memic et al. 2021). Using multiple in-season observations of multiple target variables was tested successfully for estimating crop model cultivar coefficients in a published Memic et al. (2021) study.

A flow diagram of the soil profile optimization process is shown in Fig. 3. In step 1, representative parameters for the optimization process must be carefully considered based

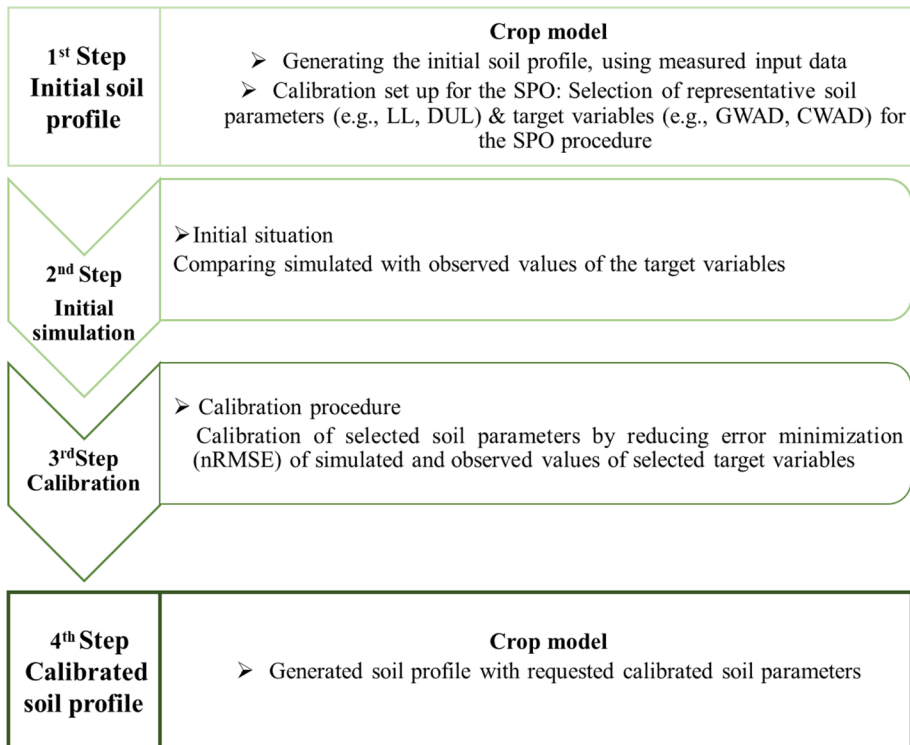


Fig. 3 Flow diagram of the soil profile optimization and simulation process with SPO

on theory, available measured data, and the study’s objective to avoid undesired autocorrelations of the optimized parameters (e.g., soil fertility and mineralization). In step 2, the DSSAT-CSM simulates crop yield based on a field-specific soil profile and evaluates it with the corresponding observed yield based on the error minimization method. In step 3, the SPO calibrates the selected soil parameters through sensitivity analysis to attain a better statistical match between simulated and observed yield. The user defines the calibration range (min/max coefficient optimization range) of the selected soil parameters by the SPO. Still, it should be done based on theory and physical soil profile logic. In the final step, the recalibrated values of the selected soil parameters are used in the crop simulation for evaluation. In the case of a successful application of the SPO, the statistical correlation of simulated and measured yield is expected to increase.

Simplified application - an illustration of the SPO working flow

The calibration process of the selected soil profile parameters is comparable to a sensitivity analysis conducted on RMSE and nRMSE calculated based on the difference between simulated and observed target variables. The idea of the optimization approach is illustrated in Fig. 4 based on an optimization example of the hydraulic parameter SLLL. To attain a better statistical fit, the value of SLLL is varied in defined min/max and increment steps of the selected coefficient for each layer to reduce the nRMSE between simulated and observed target variables while keeping already established SLLL and SDUL as primary references for the soil profile derived based on measured soil properties. Figure 4, the -y-axis shows soil layer depths in the defined soil profile, from 0 (surface) to 180 cm

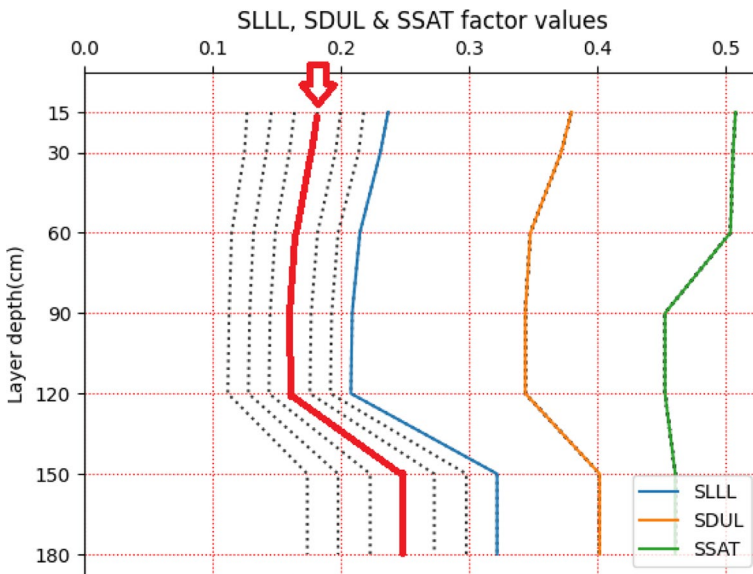


Fig. 4 Illustration of the optimization procedure of the yield-based SLLL. Optimize from the original SLLL value (SLLL through all layers) to create seven scenarios, shown as dotted lines. The lowest difference between simulated and observed is selected as “optimum” which corresponds to the SLLL blue line in this example (Color figure online)

(0–15, 15–30, 30–60, etc.). Due to simulation accuracy, simulations of the first layer are divided into 0–15 cm and 15–30 cm, as the surface layers should be at most 20 cm by the recommendation of DSSAT users. The x-axis shows SLLL, SDUL, and SSAT values for all defined layer depths. The initial SLLL values (SLLL through all layers) are the baseline (red line with an arrow pointing) for creating seven different SLLL curve scenarios, shown in Fig. 4 as dotted lines. In the example shown in Fig. 4, additional SLLL scenarios are created by reducing the original SLLL value in increments of 10%, 20%, and 30% (on the left side) of the initial SLLL line and by incrementing the initial SLLL line by 10%, 20%, and 30% (on the right side). Overall, the value setup in the SPO in Fig. 4 was ± 30 with increment steps of 10%, resulting in a total of seven SLLL curves, including the initial one. The crop model was executed for these seven SLLL scenarios, and the observed target variables (yield) were statistically analyzed. The scenario resulting in the lowest difference between simulated and observed yield was selected as “optimum”, which was the SLLL blue line scenario in this example. At the same time, SDUL and SSAT were kept constant (unchanged in the optimization process).

To clarify the mechanism of the optimization procedure, Table 2 illustrates the curve line scenario of Fig. 4 and shows the numerical influence on the SLLL values for the first layer (depth 0–15 cm).

Introduction of the SPO

In Fig. 5, the interface of the SPO is depicted by showing the eight simulation steps. Consistent with the previously discussed explanations, Fig. 5 demonstrates the interface of the SPO by optimizing the SLLL targeting GWAD with the available Riechfield dataset. To run an optimization, the user has to go through the eight simulation steps. After the SPO windows runnable is executed (1st step), all crop models available in the DSSAT shell are offered by pressing “List crop models” (2nd step). Once the desired crop growth model is selected by choosing “Load FileX/s”, all the available experiment files will be loaded into the list widget window in step 3. In the 4th step, the available treatments (TRT/s) appear for selection. In this case study, the FileX treatments correspond to the site-specific units. After selecting the site-specific units (TRT/s) in step 4, the soil profiles related to these treatments are listed (5th step). The UHIRF05001 is the soil identifier for labeling the soil profile in SOIL.SOL file. In this scenario, the soil profile remains consistent across all three years and can be optimized using three years of crop model maize parametrization and weather data. This approach allows for exploring the seasonality factor in the specific characterization of the soil profile. In step 6, multiple target variables for the optimization are based on DSSAT PlantGro.OUT or other DSSAT time-series output file (if “other” checkbox is initialized) can be selected. In step 7 (%) reduction of available parameter values),

Table 2 Numerical example of creating sensitivity analysis scenarios (1–7) for SLLL values of the first layer (depth 0–15 cm) by varying the coefficient in increment steps of 10%

	Coefficient min/max (± 30) with increment step (10%)						
Coefficients setup	–30%	–20%	–10%	Original SLLL	+10%	+20%	+30%
Calculated SLLL values optimization procedure by varying original SLLL values with the coefficient set up (1–7) for soil layer 1	0.126	0.144	0.162	0.18	0.198	0.216	0.234

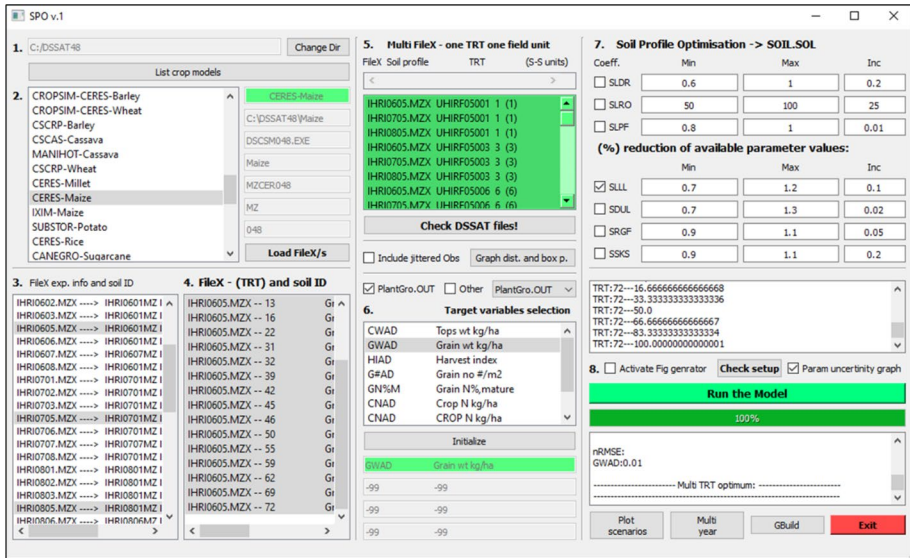


Fig. 5 Interface of the Soil Profile Optimizer with a current case study example

specific soil profile coefficients to conduct the analysis can be initialized. The coefficient initialization uses a multiplier approach, where each multiplier setup is based on the coefficient minimum/maximum values and an incremental step. This approach generates multiple soil profile scenarios. Each multiplier corresponds to percentage values, as indicated before in Table 2. In the final stage (8th step), after completing the optimization run, the user can analyze his data by entering GBuild. Moreover, the possibility is given to create a coefficient-based scenario, as shown previously in Fig. 4, by activating the Fig generator.

Four scenarios ranging from field scale to site-specific simulations

Soil sampling is labor-intensive and time-consuming. Soil sampling is often reduced to a minimum level, where few soil samples are taken randomly over a field and used for estimating field-scale soil properties. Therefore, taken samples are considered to be representative of a field-specific soil characteristic. This important implicit assumption needs to be considered because crop growth models were developed to simulate crop growth within homogeneous land units and are commonly used for evaluating the impact of management practices on yield at the field scale. In the case of a homogeneous field, the field-specific soil characterization is expected to lead to an accurate simulation of field-level yield. In cases of higher soil heterogeneity, field-scale soil profile data is not likely to capture measured spatial variability of yield caused by varying soil properties. Therefore, the number of soil samples has to be increased, and soil samples have to be taken on a site-specific (grid) level, leading to higher costs and higher costs and labor input. However, using the SPO approach for generating site-specific soil profiles attempts to overcome incorrect soil profile input information by deriving selected soil-related parameters from measured above-ground biomass data (e.g., yield, above-ground biomass, LAI, etc.). Combining soil profile optimization with the SPO using

inverse modelling resulted in a new approach, Site-Specific Optimization (SSO). The approach relies on already published findings of Batchelor et al. (2002), Batchelor and Paz (1997), Braga and Jones (2004), Thorp et al. (2008), and Wu et al. (2019).

To demonstrate the SSO's properties and function as a functional modelling approach, simulations conducted with the SSO were compared with crop model simulations of three further approaches ranging from field-scale to site-specific level. The specific characteristics of each simulation approach are shown in Table 3. Generally, a standard model approach represents the simulation on a field-scale level (field-specific soil characterization and field-specific yield), designated as Yield Simulation (YS). The DSSAT input files for this approach (YS) were created as averages of measured site-specific data and defined as field-specific level in the input files. Two additional scales, the Field-Specific Simulation (FSS, field-specific soil characterization, and site-specific yield) and the Site-Specific Simulation (SSS, site-specific soil characterization, and site-specific yield), were conducted to simulate on a site-specific level by varying the resolution of soil input information. The level of soil information required is a function of the spatial scale, with field scale simulations requiring minor soil information and SSS requiring the most detailed information. The FSS was characterized by the simulation at the site-specific scale using soil profile information averaged at the field scale. This approach was created assuming adequate soil data for site-specific simulation is unavailable, but field-level information is available. Therefore, the same field-specific soil profile of the YS was taken as crop model input data for 20 targeted site-specific units. The simulations were then compared with the site-specific observed yield measurements. In the third scenario (SSS), all spatially measured soil data were used to create a grid-specific soil profile. Further simulations were executed with the grid-specific observed yield data. The SSO approach was defined as the medium information required for the simulation using FSS and SSS data. In this approach, selected soil input parameters at the field scale were calibrated at the grid scale using grid-level observed data yield.

The general setup of the DSSAT input files was based on the simulation approach described in Table 3. The inputs regarding the primary modules, weather, management, and plant, were the same for all simulation approaches. The cultivar coefficients were available from previous work (Memic et al. 2021), based on measured yield (end-of-season) and above-ground biomass observations (three in-season observations for tops weight). In that study, the Time Series Estimator (TSE) used the observed data of the 20 grids for the experimental season 2006 for estimating cultivar coefficients. The seasons 2007 and 2008 were used for evaluation.

An overview of the soil profiles used in the four simulation approaches is indicated in Fig. 6 by showing exemplary the soil profile of grid one as it was used in the several approaches. The soil profile used in the YS and the FSS was calculated as an average soil profile from the measured texture data of the 20 grids. In the SSS for each site-specific unit, a soil profile ($n=20$) was created using measured soil data. For both approaches, an observational data file was created by adding all site-specific observations to each site-specific unit ($n=20$) for each experimental year. Besides the soil information recorded in the soil input file, the initial soil conditions were captured in the experimental file. The values for the initial soil conditions were set for all three approaches to the measured initial conditions as reported in the SSS. The soil input data in the SSO approach was created for each site-specific unit ($n=20$) by using the field-level soil profile parameters of the FSS and calibrating the parameters for each grid. The observational data file for the in-season observations was taken from the SSS.

Table 3 Characterization of four different model approaches by varying the level of input data to simulate on field-scale (Yield Simulation (YS)) to site-specific level (2. Field-Specific Simulation (FSS), 2. Site-Specific Optimization (SSO), 3. Site-Specific-Simulation (SSS)).

Simulation approach	Simulation Level	Soil input data	Model variables
Yield simulation (YS)	Field-scale	Field-specific	Field-specific observed yield and biomass data (e.g., LAI)
Field-specific simulation (FSS)	Site-specific	Field-specific	Site-specific observed yield
Site-specific optimization (SSO)	Site-specific	Generated site-specific soil profiles by calibrating field-specific soil profile	Site-specific observed yield
Site-specific simulation (SSS)	Site-specific	Site-specific based on measured values	Site-specific observed yield

```

*UHRF05001 -99 SICL 180 Field-Specific soil profile - Used for Yield Simulation (YS) and Field-Specific Simulation (FSS)
@SITE COUNTRY LAT LONG SCS FAMILY
Riech Germany -99 -99 heavy calcareous brown earth soils
@ SCOM SALB SLUI SLDR SLRO SLNLF SLFF SMHB SMPX SMKE
BN .14 6 .5 60 1 0.9 IB001 IB001 IB001
@ SLB SLMH SLLL SDUL SSAT SRGPF SSKS SBDM SLOC SLCL SLSI SLCF SLNI SLHW SLHB SCEC SADC
15 -99 .192 .363 .474 1 .15 1.32 -99 34.7 52.6 -99 -99 -99 -99 -99
30 -99 .192 .363 .474 .638 .15 1.32 -99 34.7 52.6 -99 -99 -99 -99 -99
60 -99 .211 .37 .485 .407 .15 1.29 -99 37.9 42.6 -99 -99 -99 -99 -99
90 -99 .219 .379 .481 .223 .15 1.3 -99 39.2 43.1 -99 -99 -99 -99 -99
120 -99 .22 .38 .481 .122 .15 1.3 -99 39.3 43.1 -99 -99 -99 -99 -99
150 -99 .2 .361 .485 .067 .15 1.29 -99 36 45.1 -99 -99 -99 -99 -99
180 -99 .2 .361 .485 .037 .15 1.29 -99 36 45.1 -99 -99 -99 -99 -99

*UHRF05001 -99 SICL 180 Site-specific Soil Profile based on generated Data with SPO - Site-specific Optimization (SSO)
@SITE COUNTRY LAT LONG SCS FAMILY
Riech Germany -99 -99 heavy calcareous brown earth soils
@ SCOM SALB SLUI SLDR SLRO SLNLF SLFF SMHB SMPX SMKE
BN .14 6 0.5 80.0 1 0.9 IB001 IB001 IB001
@ SLB SLMH SLLL SDUL SSAT SRGPF SSKS SBDM SLOC SLCL SLSI SLCF SLNI SLHW SLHB SCEC SADC
15 -99 0.202 .363 .474 1 .15 1.32 -99 34.7 52.6 -99 -99 -99 -99 -99
30 -99 0.202 .363 .474 0.67 .15 1.32 -99 34.7 52.6 -99 -99 -99 -99 -99
60 -99 0.222 .37 .485 0.427 .15 1.29 -99 37.9 42.6 -99 -99 -99 -99 -99
90 -99 0.23 .379 .481 0.234 .15 1.3 -99 39.2 43.1 -99 -99 -99 -99 -99
120 -99 0.231 .38 .481 0.128 .15 1.3 -99 39.3 43.1 -99 -99 -99 -99 -99
150 -99 0.21 .361 .485 0.07 .15 1.29 -99 36 45.1 -99 -99 -99 -99 -99
180 -99 0.21 .361 .485 0.039 .15 1.29 -99 36 45.1 -99 -99 -99 -99 -99

*THRIO40101 IBSNAT SIL 180 Site-specific Soil Profile based on measured data - Site-specific simulation (SSS)
@SITE COUNTRY LAT LONG SCS FAMILY
Riech Germany -99 -99 unknown
@ SCOM SALB SLUI SLDR SLRO SLNLF SLFF SMHB SMPX SMKE
-99 .14 6 .5 60 1 1 SA001 SA001 SA001
@ SLB SLMH SLLL SDUL SSAT SRGPF SSKS SBDM SLOC SLCL SLSI SLCF SLNI SLHW SLHB SCEC SADC
15 -99 .114 .296 .527 1 .68 1.16 1.16 11.3 57.9 -99 -99 -99 -99 -99
30 -99 .11 .288 .52 .638 .68 1.18 1 11.3 57.9 -99 -99 -99 -99 -99
60 -99 .098 .211 .436 .407 2.59 1.42 .68 10.7 29.1 -99 -99 -99 -99 -99
90 -99 .141 .282 .457 .223 1.32 1.37 .26 21.5 45.3 -99 -99 -99 -99 -99
120 -99 .14 .281 .457 .122 1.32 1.37 .25 21.5 45.3 -99 -99 -99 -99 -99
150 -99 .244 .406 .45 .067 .09 1.39 .2 42.5 49.9 -99 -99 -99 -99 -99
180 -99 .244 .406 .45 .037 .09 1.39 .2 42.5 49.9 -99 -99 -99 -99 -99

```

Fig. 6 Soil profile input files for the exemplary chosen grid 1 for the four different simulation approaches I. Yield Simulation (YS) and II. Field-Specific Simulation (FSS), III. Site-Specific Optimization (SSO), IV. Site-Specific Simulation (SSS)

The SPO was developed as a software plug-in and can be downloaded as freeware from the GitHub account (<https://github.com/memicemir>).

Statistical evaluation of modelling results

The correlation-regression-based statistical method in this study is based on the coefficient of determination (R^2). R^2 is a statistical measure of how well data fits a regression line. The linear model, which correlates simulated and observed data, is shown in Eq. 5 (Willmott, 1981; Yang et al. 2014).

$$y = \alpha + \beta x + \epsilon \quad (5)$$

R^2 indicates the strength of the linear relationship: $R^2 = 1$ indicates a perfect fit, while $R^2 = 0$ indicates no linear relation. The R-squared statistic only captures linear associations and not variations in the fit relative to actual data (Willmott, 1981; Yang et al. 2014).

To overcome the limitations of correlation-based statistics, efficiency measures have been developed to assess deviations ($d = y - x$) directly. A statistical index is the mean error (E) where $i = 1, 2, \dots, n$ (Eq. 6) (Addiscott & Whitmore, 1987; Yang et al. 2014).

$$E = \sum (y_i - x_i) / n \quad (6)$$

The mean error indicates whether the model underestimates ($E < 0$) or overestimates ($E > 0$) the observed data. However, a drawback of E is that positive and negative errors can offset each other, resulting in $E = 0$ (Yang et al. 2014). New methods that rely on the sum of squares were introduced to handle this limitation. This study examines the use of

root mean square error (RMSE) shown in Eq. (7) and the modeling efficiency (EF) shown in Eq. (8).

$$RMSE = \sqrt{\sum (y_i - x_i)^2 / n} \quad (7)$$

The root mean square error (RMSE) is commonly used in model calibration and validation to measure the deviation ($y-x$) between predicted and observed values. Its unit of measure is the same as the deviation (Loague & Green, 1991).

$$EF = 1 - \sum (y_i - x_i)^2 / \sum (x_i - \bar{x})^2 \quad (8)$$

EF ranges from -1 to 1 and was introduced by Nash and Sutcliffe (1970) to evaluate river flow models. The metric of the EF is dimensionless. An $EF=1$ indicates that the model's output perfectly matches the observed data. If the $EF < 1$, it means that the simulation is realistic but not perfect. When the EF value is less < 0 , it means that the model's predictions are worse than just using the observed mean (\bar{x}) instead of the modeled values (y_i). Generally, an EF value greater than 0 is an essential criterion for determining the goodness of fit between the simulated and observed data. EF has been used widely in model evaluation and has been called by various names such as Nash-Sutcliffe Efficiency (NSE), coefficient of efficiency, and modeling efficiency (Loague & Green, 1991; Yang et al. 2014).

Results

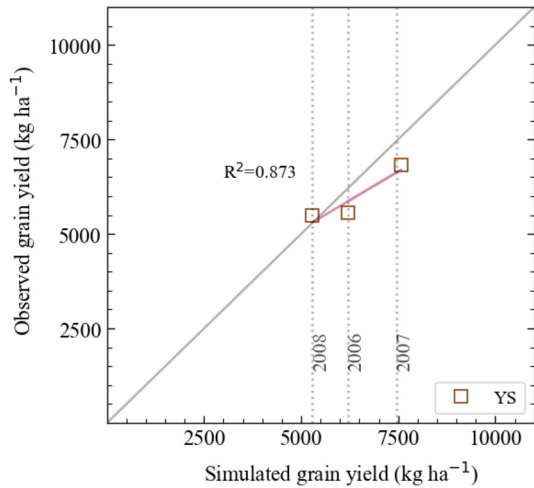
Yield simulation (YS) on a field scale level

In 2006, the field-specific observed grain yield of the Riech field was 5549 kg ha^{-1} ; in 2007, 6824 kg ha^{-1} ; and in 2008, 5485 kg ha^{-1} . The yield distribution was strongly affected by seasonality because of varying weather conditions and can be associated with temporal variability. The field-specific grain yield was simulated in the approach named YS. The results are shown in Fig. 7 as simulated vs. observed field-specific grain yield for the experimental years 2006, 2007, and 2008, with a corresponding R^2 of 0.87 . Simulated and observed GWAD with the corresponding RMSE is shown in Table 4. The smallest RMSE (201 kg ha^{-1}) was observed in 2008. In 2006, the RMSE of simulated vs. observed GWAD was 649 kg ha^{-1} and 757 kg ha^{-1} in 2007. The RMSE of the field-scale level simulation was approximately 10% of the measured yield and thus within the usual acceptable range of model studies.

Table 4 Results of the Yield Simulation (YS) Field-specific observed and simulated grain yield (GWAD) and RMSE for the Riech field (2006–2008)

Year	Observed GWAD kg/ha ⁻¹	Simulated GWAD	RMSE
2006	5549	6198	649
2007	6824	7581	757
2008	5485	5284	201

Fig. 7 Simulated vs. observed field-specific grain yield (kg ha^{-1}) for the experimental years 2006–2008, based on the Yield Simulation (YS)



Site-specific optimization (sso) over multiple years - comparison of the sso with different simulation approaches

A comparison of the four simulation approaches is shown in Fig. 8 by illustrating the simulated vs. observed grain yield and the related R^2 for the YS and the FSS (Fig. 8a), the SSS (Fig. 8b), and the SSO (Fig. 8c). In Fig. 8a, the simulation results of the FSS were illustrated together with the simulation results of the YS. The simulated grid-specific grain yield of the FSS resulted in the exact simulated yield for each grid and year. Comparing the simulated values of the FSS with the simulation results of the YS showed that the simulated data points can be assigned to the three simulated values of the YS. Therefore, the simulated values of the FSS reflect temporal variability for each experimental year (2006–2008), but merely the observed site-specific grain yield was depicted on the y-axis with all in-season variability. The minor variations in the simulated yield (Fig. 8a, x-axis) were due to measured soil initial conditions included in the simulation procedure

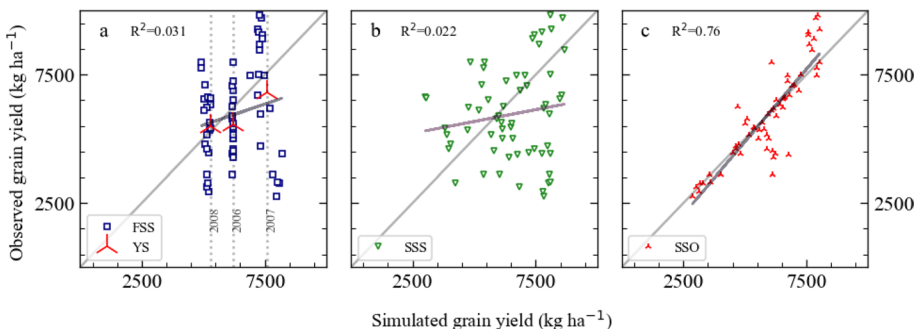


Fig. 8 Comparison of four different simulation approaches by showing simulated vs. observed grain yield (kg ha^{-1}) for the Field-Specific Simulation (FSS) combined with the Yield Simulation (YS) **a** the Site-Specific Simulation (SSS) **b** and the Site-Specific Optimization (SSO) conducted with the three selected soil profile parameters (SLRO, SLLL, SRGF) by targeting GWAD over three years **c**. All results are illustrated for 2006–2008

for all three years site-specifically. The RMSE calculated over three years and all 20 grids was approximately 25%. The FSS could not explain site-specific yield variability based on field-specific soil characterization, as indicated by the low R^2 value of 0.03 and EF value of -0.13 . Figure 8b shows the results of the SSS by comparing the simulated site-specific yield (x-axis) with the site-specific measured yield (y-axis). The SSS, characterized by the simulation approach of simulating with site-specific created soil profiles based on measured soil texture values, led to an R^2 of 0.02 and an EF of -0.5 . Besides the low R^2 , the SSS resulted in an overall RMSE over three years (2006–2008) and all 20 grids of around 30%. The SSS approach could not explain site-specific yield variability over three experimental seasons based on site-specific measured soil texture and pedo-transfer functions.

The results of the SSO, based on 20 site-specific soil profiles generated with the SPO by minimizing error in simulated and observed grain yield over three years (multiple years), are shown in Fig. 8c. The simulation conducted with the site-specific optimized soil profiles resulted in an R^2 of 0.76 and an EF of 0.75. Out of all the approaches tested, the SSO showed the best statistical fit compared to the FSS and SSS. In this crop model-based analysis, 76% of the site-specific yield variability (yield = dependent variable) was explained by targeted independent variables (soil profile parameters). The comparison of R^2 of the three simulation approaches on a site-specific level showed that the SSO gave the best result (Fig. 8c). The SSO was conducted with the measured soil initial conditions taken from the SSS. Even higher correlations were achieved by running the SSO with varying initial soil conditions. Setting the initial soil conditions on averaged values led to an R^2 of 0.81, while setting them to 0 resulted in an R^2 of 0.79.

Comparison of the soil profile parameters in the optimization process

The site-specific soil profile parameters are selected based on optimizing multiple choice parameters to establish the best-performing parameter combination to explain in-field yield variability. Since the experimental period was over three years, calibrating a maximum of three parameters was expected to deliver mathematically meaningful results. The influence of each soil profile input parameter available in the SPO was tested for the Riech field data set in three different optimization scenarios, and the results are shown in Table 5. The optimization setups are based on a multi-year approach. Therefore, the optimum was calculated for 20 grids over three years. The optimization range for the 1st optimization scenario was set to $\pm 10\%$ of the original values. For the 2nd optimization scenario, the optimization range was $\pm 20\%$, and the 3rd $\pm 30\%$ of the initial values.

The optimization of SDUL looked promising and led to significant improvements in the first two optimization steps. However, the 3rd optimization step could not be conducted because the SDUL values exceeded the SAT values, which resulted in a model abort. For the optimization of SLLL, a relevant improvement was observed in the 2nd step. In the 3rd step, R^2 increased to 0.39, while the RMSE also increased. Of all optimized soil profile parameters, the 2nd optimization step of the SLLL resulted in the lowest RMSE (1078 kg ha^{-1}), which justified the selection for the optimum scenario. The optimization of the SLRO showed the highest statistical impact on optimizing soil profile parameters. Compared to all other soil profile parameters, R^2 in the 2nd optimization step was already 0.44. The best result for the SLRO optimization was found for the optimization range from 60 to 100, leading to an R^2 of 0.48. The SRGF showed the best statistical fit in the 3rd optimization step, which resulted in R^2 of 0.35 and RMSE of 1201 kg ha^{-1} . As the approach focused on optimizing the soil profile parameters related to the soil water balance, SLPF

Table 5 Comparison of the R² and the RMSE (kg ha⁻¹) (RMSE calculated as mean over 20 grids) between simulated and observed grain weight for three optimization scenarios for 20 grids over three years (2006–2008), by varying the optimizations ranges of selected soil profile parameter (varying degrees in 10% steps)

Soil Profile Parameter available in the SPO	1st optimization scenario ($\pm 10\%$ of the original values)		2nd optimization scenario ($\pm 20\%$ of the original values)		3rd optimization scenario ($\pm 30\%$ of the original values)		4th optimum scenario (three selected parameters, with specific Optimization Range OR) Site-Specific Optimization (SSO)			
	R ² *	RMSE**	R ² *	RMSE**	R ² *	RMSE**	OR	R ² *	RMSE**	PD***
I. SDUL Drained upper limit (cm ³ cm ⁻³)	0.28	1122	0.38	1122	–	–	–	–	–	–
II. SLLL Lower Limit of plant extractable water (cm ³ cm ⁻³)	0.28	1106	0.37	1078	0.39	1086	$\pm 20\%$	0.34	1078	18%
III. SLRO Runoff curve number (unitless)	0.27	1417	0.44	1164	0.45	1148	60–100	0.48	1098	18%
IV. SLPF Fertilization factor (0–1 scale)	0.13	1362	0.30	1197	0.36	1148	–	–	–	–
V. SRGF Root growth factor (0–1 scale)	0.14	1365	0.30	1228	0.35	1201	$\pm 30\%$	0.35	1201	20%
VI. SSKS Soil saturated hydraulic conductivity (0.05 cm h ⁻¹ for clayey soils, - to 63 cm h ⁻¹ for sandy soils)	0.03	1533	0.03	1532	0.03	1533	–	–	–	–
VII. SLDR Drainage rate (Fraction day ⁻¹)	0.02	1531	0.03	1517	0.04	1495	–	–	–	–
The selected parameter in the optimum procedure (SLLL, SLRO & SRGF)	0.56	909	0.66	753	0.74	611	–	0.76	600	10%

Bold was used to highlight the results of the optimum procedure

*R² of simulated vs. observed grain weight ** Average RMSE (kg ha⁻¹) of simulated and observed grain weight ***Percent Difference PD (%)

was not considered site-specific. Optimization of drainage rate (SLDR) and saturated hydraulic conductivity (SSKS) did not change for different scenarios and did not lead to relevant improvements (results not shown).

Out of the analysis of the soil profile parameters shown in Table 5, SLLL ($\pm 20\%$), SLRO (60–100), and SRGF ($\pm 30\%$) were selected as relevant soil profile parameters for the optimum scenario. The defined optimum scenario was the same as the SSO scenario shown in Fig. 8, leading to an R^2 of 0.76. Besides the numerical illustration in Table 5, the impact of optimizing a single soil profile parameter in the optimum scenario is illustrated in Fig. 9. Observed vs. simulated grain yield for 20 grids over three years for the optimization of one single soil parameter (SLLL (Fig. 9b), SRGF (Fig. 9c), SLRO (Fig. 9d)) were compared with the FSS (Fig. 9a) and the SSO (Fig. 9e). The optimization of SLLL, SRGF, and SLRO showed a similar pattern of outliers. By comparing the RMSE of each grid for the optimization with SLLL, SRGF, and the SLRO over the three experimental years, the same grids with an exceptionally high RMSE appeared (data not shown). However, the grids with high RMSE varied yearly, and no consistent pattern was observed. Overall, the combination of the three selected soil profile parameters in the SSO reduced the mean RMSE to 600 kg ha^{-1} .

Influence of the soil optimization procedure on critical processes of the soil water balance

Table 6 compares the simulated values of two simulation approaches, SSS and SSO, presenting the impact of the soil profile optimization on critical processes related to the soil water balance. The simulated values were calculated yearly as means over 20 grids. The values are presented in absolute numbers (mm), and in addition, the percentage change (%) of the SSO related to the SSS is provided (Table 6).

Cumulative precipitation was calculated from sowing to harvest in 2006 as 425 mm (181 simulated days), in 2007 as 386 mm (177 simulated days), and in 2008 as 439 mm (169 simulated days). In all three years, water stress was detected (data not shown). Even though precipitation was lowest in 2007, the simulated water stress factors were higher in 2006 and 2008.

To compare the evapotranspiration and their referring processes, soil evaporation, and transpiration between the simulation approaches SSS and SSO, the values were shown as cumulative values from the day of sowing until the day of harvest. The percentage change of the cumulative transpiration of the SSO showed an increase of 6% in 2006, 1% in 2008, and a decrease of 5% in 2007. The percentage change between SSS and SSO for the

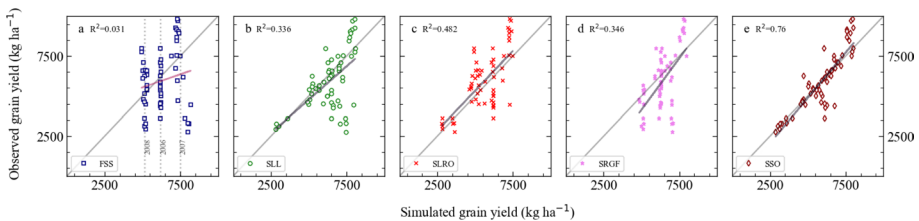


Fig. 9 Comparison of simulated vs. observed grain yield (kg ha^{-1}) for the Field-specific simulation (FSS) **a** SLLL (Lower limit) **b** SLRO (Runoff curve number) **c** SRGF (Root growth factor) **d** Site-specific optimization SSO **e** for 20 grids optimized by targeting GWAD over three years 2006–2008

Table 6 Comparison of the simulated values (calculated as mean over 20 grids /given as absolute values (mm) and the percentage change (%) of Site-Specific Optimization (SSO) related to the Site-Specific Simulation (SSS)) for the processes evapotranspiration, transpiration, soil evaporation, extractable water for the years 2006, 2007 and 2008

Variables of the soil water balance of DSSAT	Simulation approach	Simulated values calculated as mean over 20 grids								
		2006			2007			2008		
		DAP*	mm	%**	DAP*	mm	%**	DAP*	mm	%**
Cumulative precipitation	–	181	425	–	177	386	–	169	439	–
Cumulative evapotran-spiration	SSS	181	480	–	177	515	–	169	515	–
	SSO	181	466	–5	177	511	–1	169	475	–8
Cumulative transpiration	SSS	181	302	–	177	343	–	169	299	–
	SSO	181	319	+6	177	327	–5	169	303	+1
Cumulative soil evaporation	SSS	181	178	–	177	172	–	169	216	–
	SSO	181	147	–17	177	147	–15	169	171	–20
Extractable water	SSS	0	271	–	0	240	–	0	255	–
	SSO	0	300	+11	0	281	+17	0	287	+12
	SSS	181	204	–	177	106		169	159	
	SSO	181	171	–16	177	120	+13	169	150	–6

cumulative soil evaporation showed 17% (2006), 15% (2007), and 20% (2008) a consistent reduction. Overall, the SSO decreased cumulative evapotranspiration over all three years (–5% in 2006, –1% in 2007, –8% in 2008).

To show the effect of soil profile optimization on plant available water, the day of sowing and the day of harvest were picked. In the SSO approach, plant available water was at the day of sowing, under water-saturated conditions in all three years, consistently higher (2006, 11%, 2007 17%, and 2008 12%). At harvest day, the SSO resulted in 2006, respectively 2008 in a decrease of 16% and 6%, and an increase of 13% in 2007. Moreover, the extractable water was 106 mm (SSS) and 120 mm (SSO), the lowest in 2007, when cumulative precipitation was also observed.

In conclusion, the impact of the soil profile optimization on the selected processes was higher on directly linked soil water processes like soil evaporation and plant available water. It can be assumed that the consistent decrease of evapotranspiration in the SSO is due to reduced soil evaporation. Comparing soil available water on the day of harvest with the transpiration indicated similar trends. The higher cumulative transpiration in 2006 and 2008 of the SSO leads to the assumption that the plant water uptake in this approach was higher and could explain why the extractable water at harvest day was lower. In 2007, a reversed trend was observed.

Discussion

The simulation approaches conducted in this study demonstrated the challenges of site-specific crop modelling. The YS scenario showed an accurate model performance at the field scale level with an R^2 of 0.87 and RMSE of approximately 500 kg ha⁻¹. Compared to YS, FSS, and the SSS conducted at the site-specific level could have led to an adequate simulation result. FSS showed that a general soil profile (field-specific) did not sufficiently explain

yield variability on a site-specific level. Only slight variations due to varying initial conditions stored in the experimental file were captured in the simulated values; several studies reported that precise soil input data is needed for a successful simulation on a site-specific level (Batchelor et al. 2002; Braga & Jones, 2004). However, additional site-specific soil input data does not automatically lead to an acceptable simulation result, as shown in the SSS conducted for the Riech field data. Even though soil input data was available on the site-specific level, the SSS in this study only gave an R^2 of 0.02, an EF of -0.51 , and an RMSE of 1800 kg ha^{-1} . Based on the assumption of Maestrini and Basso (2018) that spatial variability is caused by soil variability, an explanation could be found in the underlying structure of the DSSAT crop model. Initially, the DSSAT model was designed to simulate yield on an area basis (field scale level) under the assumption of land unit homogeneity. Model parameters must be downscaled to simulate yield on a site-specific level, as Pasquel et al. (2022) reported. Only limited research has yet to be done to develop downscaling methods. However, these methods are crucial in determining the size of meaningful site-specific units, especially from the perspective of crop modelling. Link et al. (2006) indicated that at a certain resolution of the grid size, no further model improvement could be reached because the simulations cannot address the temporal variability across seasons. Conversely, larger grids could describe temporal variability but not spatial yield variability.

The current application of the DSSAT model on a site-specific level points to the challenge of using a one-dimensional soil characterization in a PA approach (Sadler et al. 2000). The functionality of the one-dimensional soil model in DSSAT inherently considers only horizontal processes (Ritchie, 1998). Sadler et al. (2000) reported in their study that possible interactions of horizontal water transfer via runoff or flow are rarely considered. For the SSS with the Riech field data set, it can be assumed that the measured soil samples needed to be more appropriate, and more data is required to generate suitable soil profiles.

Batchelor et al. (2004); Thorp et al. (2008) reported that a systematic soil profile optimization tool for site-specific modelling with the DSSAT model is needed. However, the last version of the Apollo Decision Support System (DSS) was available for DSSAT 3.5. Due to model changes, some of the optimization parameters in Apollo are unavailable in the current DSSAT version. As a new systematic approach, the SPO was designed for the current DSSAT version 4.8 based on a generic soil profile structure and can be used with future versions of DSSAT. The SPO calibrated for SLLL, SLRO, and SRGF of grid-level soil profiles by targeting the grain weight on a site-specific level gave an R^2 of 0.76, an EF of 0.75, and a mean RMSE of 600 kg ha^{-1} for the Riech field data.

Statistical analysis is a crucial component in calibrating and evaluating crop growth models. It helps to ensure the accuracy and reliability of the models, as discussed in Yang et al. (2014). To conduct a representative statistical conclusion, several statistics are needed. The SPO is primarily based on the nRMSE minimization method due to the advantage of optimizing specific parameters based on multiple target variables with different units. In this study, the R^2 and RMSE were primarily utilized for model evaluation. The EF was calculated for the solutions obtained by the SPO as an additional statistical evaluation of the optimized model performance. Dimensionless statistics such as EF are widely used in model evaluation and can potentially increase the SPO's accuracy. Due to the goal of this study, only soil profile parameters with an influence on the PAWC are considered. The PAWC, a main driver of yield variability in maize production, is driven by the hydraulic parameters calculated with the PTF. As shown in Sect. 2.2, the calculation of the PAWC in the DSSAT model is driven by SDUL and SLLL, determined in the approach of Saxton et al. (1986). Several studies reported the challenges of using PTFs in modelling, reported difficulties in their application, and recommended optimization (Gijsman et al.

2002; Patil & Singh, 2016; van Looy et al. 2017). Therefore, this study focused on the optimization of SDUL and SLLL. Due to model abortions during the optimization of SDUL, especially in optimization scenarios of higher ranges, SDUL was rejected for the optimization procedure. Compared to the SDUL, the SLLL showed higher robustness in the optimization procedure. However, Wu et al. (2019) reported similar effects on SDUL and SLLL simulation results in their optimization approach. Moreover, errors in the estimation of SLLL have a more significant impact on the simulation because the SLLL applies to much smaller values than the DUL (Gijsman et al. 2002). Patil and Singh (2016) reported a general RMSE of PTFs to predict soil water retention from 0.007 to 0.07 m³ m⁻³. The selected optimization range of 20% for the optimum scenario of the SLLL is in an acceptable range. Ritchie (1998) pointed out that the approach to calculating the infiltration in DSSAT, the curve number method, needs to be revised. Similar results were reported by Sadler et al. (2000). Due to the assumption that the approach of the curve number method and the one-dimensional model setup lead to model inaccuracy, the SLRO 60–100 optimization range is appropriate. It can be assumed that the optimization of SLLL influences the PAWC and SLRO in the daily simulated water content. Usually, optimizing hydraulic parameters is coupled with calibrating the root growth parameter (Thorp et al. 2008; Wu et al. 2019). Overall, the selected parameters led to good results, and the influences were similar to those reported by Wu et al. (2019).

The comparison of simulation results between SSS and SSO for evapotranspiration, transpiration, soil evaporation, and extractable water showed the optimization procedure's impact on these general processes. Boote et al. (2008) reported the importance of adequately estimating the soil water-holding parameters and root growth to receive a satisfactory tipping bucket soil water balance model output. The role of a precise soil water balance on crop water stress signals is described. Depending on the simulated water stress factor, the model alters crop assimilation, expansive growth processes, and most crop phenological progressions (Boote 2008). These findings showed the dimension of the conducted SSO, and it can be assumed that through this optimization also, growth processes were modified. Precise yield predictions in site-specific modelling are essential for site-specific crop management. Only accurate model input parameters can lead to model simulations, which can be further used for model testing to determine agronomic and economic outcomes of specific management practices (Braga & Jones, 2004). Based on the findings of this study, a tactical approach for a variable rate input application in crop production for farmers can be set up (Maestrini & Basso, 2018). As Memic et al. (2019) indicated, using crop models as a decision support tool for site-specific N application rates also needs to capture soil parameters, representing the spatial and temporal variability.

Conclusion

The need for crop models in PA as a tool to develop risk management is undisputed. However, model simulations depend on the quality of input data. In general, soil sampling is time-consuming, labor-intensive, and depending on the level of field heterogeneity, often inappropriate for site-specific management. An inverse modelling approach was developed and built on the assumption that yield patterns can lead to insights into given variability of soil properties. The SPO software developed in this study can be used as an external plugin program for the current DSSAT version 4.8. The case study for a 3-year field trial with

maize showed the possibility of calibrating soil profile parameters from a general soil profile. A strong influence of soil profile parameters affecting PAWC was shown.

The SPO as an external plug-in program for the DSSAT model led to promising results for the current data set and improved final yield simulations. Overall, the SPO could be a valuable methodology for a generic optimization of soil profiles for the DSSAT model. However, further testing and verification need to be conducted with independent data sets. Furthermore, combining the SPO with site-specific fertilizer optimization would be an envisioned additional step.

Author contributions All authors contributed to the study's conception and design. JT performed material preparation and analysis. JT wrote the first draft of the manuscript, and all authors commented on previous versions. All authors, JT, EM, WDB, and SG read and approved the final manuscript. Conceptualization: JT, Methodology: JT, EM, Formal analysis and investigation: JT, Writing—original draft preparation: JT; Writing—review and editing: JT, EM, WDB, SGH, Funding acquisition: SGH, Resources: SGH, Supervision: SGH.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors would like to acknowledge the research project CHARGE (Field investigation on the impact of high voltage direct current (HVDC) underground cables on soils and crop production) that deals with PTKA/KIT (L7521101). The funding for this project by the Ministry of the Environment, Climate Protection, and the Energy Sector Baden-Württemberg is gratefully acknowledged. The project DiWenkLa (Digital Value Chains for a Sustainable Small-Scale Agriculture) is supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany. The Federal Office for Agriculture and Food (BLE) coordinates support for digitalisation in agriculture as a funding organization; grant number 28DE106A18 is gratefully acknowledged. DiWenkLa is also supported by the Ministry for Food, Rural Areas, and Consumer Protection Baden-Württemberg.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbaspour, K., Kasteel, R., & Schulin, R. (2000). Inverse parameter estimation in a layered unsaturated field soil. *Soil Science*, *165*(2), 109–123.
- Adams, W. A. (1973). The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. *Journal of Soil Science*, *24*(1), 10–17. <https://doi.org/10.1111/j.1365-2389.1973.tb00737.x>.
- Adamchuk, V. I., Gebbers, R., & . (2010). Precision agriculture and food security. *Science*, *327*, 828–831.
- Addiscott, T. M., & Whitmore, A. P. (1987). Computer simulation of changes in soil mineral nitrogen and crop nitrogen during autumn, winter and spring. *The Journal of Agricultural Science*, *109*(1), 141–157. <https://doi.org/10.1017/S0021859600081089>.
- Aitkenhead, M., Amelung, W., Assouline, S., Baveye, P., Berli, M., Brüggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Hopmans, J. W., Javaux, M., Or, D., Roose, T., Schnepf, A., Vanderborght, J., Vereecken, H., Young, M., & H., & Young, I. M. (2016). Modeling soil processes: Review, Key challenges, and New perspectives. *Vadose Zone Journal*. <https://doi.org/10.2136/vzj2015.09.0131>

- Hoogenboom, G., C.H. Porter, K.J. Boote, V. Shelia, P.W. Wilkens, U. Singh, J.W. White, S. Asseng, J.I. Lizaso, L.P. Moreno, W. Pavan, R. Ogoshi, L.A. Hunt, G.Y. Tsuji, and J.W. Jones (2019). The DSSAT crop modeling ecosystem. In K. Boote (Ed.), *Advances in crop modelling for a sustainable agriculture* (pp. 173–216). Burleigh Dodds Science Publishing. <https://doi.org/10.19103/AS.2019.0061.10>
- Basso, B., Dobrowolski, J., & McKay, C. (2017). From the Dust Bowl to drones to Big Data: The Next Revolution in Agriculture. *Georgetown Journal of International Affairs*, 18(3), 158–165. <https://doi.org/10.1353/gia.2017.0048>.
- Batchelor, W. D., Basso, B., & Paz, J. O. (2002). Examples of strategies to analyze spatial and temporal yield variability using crop models. *European Journal of Agronomy*, 18(1–2), 141–158. [https://doi.org/10.1016/S1161-0301\(02\)00101-6](https://doi.org/10.1016/S1161-0301(02)00101-6).
- Batchelor, W. D., Paz, J. O., & Thorp, K. R. (2004). Development and evaluation of a decision support system for precision agriculture Proceedings of the 7th International Conference on Precision Agriculture, ASA-CSSA-SSSA, Madison, WI, U.S.A.
- Baumer, O. W., & Rice, J. W. (1988). Methods to predict soil input data for DRAINMOD (88-2564). American Society of Agricultural Engineers.
- Bernardo, Maestrini Bruno, Basso (2018) Drivers of within-field spatial and temporal variability of crop yield across the US Midwest Abstract Scientific Reports, 8(1). <https://doi.org/10.1038/s41598-018-32779-3>
- Boote, K. J. (2019). Advances in crop modelling for a sustainable agriculture. *Burleigh Dodds Science Publishing*. <https://doi.org/10.1201/9780429266591>.
- Boote, K. J., Hoogenboom, G., Jones, J. W., & Wilkerson, G. (1997). Evaluation of the CROPGRO-Soybean model over a wide range of experiments. In M. J. Kropff, P. S. Teng, P. K. Aggarwal, J. Bouma, B. A. M. Bouman, J. W. Jones, & H. H. van Laar (Eds.), *Systems Approaches for Sustainable Agricultural Development*. Springer.
- Boote, K. J., Hoogenboom, G., & Jones, & J. W., Sau, F.,. (2008). Experience with water balance, evapotranspiration, and prediction of water stress effects in the CROPGRO model. In L. R. Ahuja, V. R. Reddy, S. A. Saseendran, & Q. Yu (Eds.), *Advances in agricultural systems modeling* (pp. 59–103). Wiley.
- Bouma, J. (1992). Using Soil Survey Data for Quantitative Land Evaluation. In B. A. Stewart & R. Lal (Eds.), *Advances in Soil Science Soil Restoration* (pp. 177–213). Springer.
- Braga, R., & Jones, J. W. (2004). Using optimization to estimate soil inputs of crop models for use in site-specific management. *Transactions of the ASAE*, 47(5), 1821–1831. <https://doi.org/10.13031/2013.17599>.
- Campbell, G. S. (1974). A simple method for determining unsaturated conductivity from moisture retention data. *Soil Science*, 117, 311–314.
- Chen, X., Jiao, X., Lü, H., Salahou, M. K., & Zhang, Y. (2022). Inverse Modelling to Estimate Soil Hydraulic Properties at the Field Scale. *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2022/4544446>
- Claupain, W., Graeff, S., & Link, J. (2013). Comparison of uniform control and site-specific model-based nitrogen prescription in terms of grain yield, nitrogen use efficiency and economic aspects in a heterogeneous corn field. *Focus on plant cultivation sciences*, 65, 6.
- Dalglish, N. P., & Foale, M. A. (1998). Soil matters: Monitoring Soil Water and nutrients in Dryland Farming. CSIRO Tropical Agriculture, Agricultural Production Systems Research Unit.
- Gijsman, A., Jagtap, S., & Jones, J. (2002). Wading through a swamp of complete confusion: How to choose a method for estimating soil water retention parameters for crop models. *European Journal of Agronomy*, 18(1–2), 77–106. [https://doi.org/10.1016/S1161-0301\(02\)00098-9](https://doi.org/10.1016/S1161-0301(02)00098-9).
- Gijsman, A. J., Thornton, P. K., & Hoogenboom, G. (2007). Using the WISE database to parameterize soil inputs for crop simulation models. *Computers and Electronics in Agriculture*, 56(2), 85–100. <https://doi.org/10.1016/j.compag.2007.01.001>.
- Green, R. E., & Loague, K. (1991). Statistical and graphical methods for evaluating solute transport models overview and application. *Journal of contaminant hydrology*. [https://doi.org/10.1016/0169-7722\(91\)90038-3](https://doi.org/10.1016/0169-7722(91)90038-3)
- Godwin, D., Jones, C., Ritchie, J., Vlek, P., & Youngdahl, L. J. (1984). The water and nitrogen components of the CERES models (Proc. Int. Symp. on Minimum Data Sets for Agrotechnology Transfer.). Prades, A., pp. 101–106.
- Hartmann, K., Krois, J., & Waske, B. (2018). E-Learning Project SOGA: Statistics and Geospatial Data Analysis. Department of Earth Sciences, Freie Universitaet Berlin.
- Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheny, E., Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B. T., Kersebaum, K. C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., & Ewert, F. (2016). Impact of spatial soil and climate

- Input Data Aggregation on Regional Yield simulations. *PLoS One*, 11(4), e0151782. <https://doi.org/10.1371/journal.pone.0151782>.
- Indoria, A. K., Reddy, K., & S., & Sharma, K. L., (2020). Hydraulic properties of soil under warming climate. *Climate Change and Soil Interactions* (pp. 473–508). Elsevier.
- Kamali, H. R., & Zand-Parisa, S. (2016). Optimization of a new inverse method for estimation of individual soil hydraulic parameters under field conditions. *Transactions of the ASABE*, 59(5), 1257–1266. <https://doi.org/10.13031/trans.59.11414>.
- Link, J., Graeff, S., Batchelor, W. D., & Claupein, W. (2006). Spatial variability and temporal stability of corn (*Zea mays* L.) grain yields – relevance of grid size. *Archives of Agronomy and Soil Science*, 52(4), 427–439. <https://doi.org/10.1080/03650340600775487>.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431–441. <https://doi.org/10.1137/0111030>.
- Memic, E., Graeff, S., Claupein, W., & Batchelor, W. D. (2019). GIS-based spatial nitrogen management model for maize: Short- and long-term marginal net return maximising nitrogen application rates. *Precision Agriculture*, 20(2), 295–312. <https://doi.org/10.1007/s11119-018-9603-4>.
- Memic, E., Graeff, S., Boote, K. J., Hensel, O., & Hoogenboom, G. (2021). Cultivar Coefficient Estimator for the Cropping System Model based on Time-Series Data: A case study for soybean. *Transactions of the ASABE*, 64(4), 1391–1402. <https://doi.org/10.13031/trans.14432>.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., & Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture*, 23(4), 1489–1513. <https://doi.org/10.1007/s11119-022-09885-4>.
- Patil, N. G., & Singh, S. K. (2016). Pedotransfer functions form estimating Soil Hydraulic properties. *A Review Pedosphere*, 26(4), 417–430. [https://doi.org/10.1016/S1002-0160\(15\)60054-6](https://doi.org/10.1016/S1002-0160(15)60054-6).
- Paz, J., Batchelor, W. D., Babcock, B., Colvin, T., Logsdon, S., Kaspar, T., & Karlen, D. (1999). Model-based technique to determine variable rate nitrogen for corn. *Agricultural Systems*, 61(1), 69–75. [https://doi.org/10.1016/S0308-521X\(99\)00035-9](https://doi.org/10.1016/S0308-521X(99)00035-9).
- Rawls, W. J., Brakensiek, D. L., & Saxton, K. E. (1982). Estimation of Soil Water properties. *Transactions of the ASAE*, 25(5), 1316–1320. <https://doi.org/10.13031/2013.33720>.
- Ritchie, J. T. (1985). A User-Orientated Model of the Soil Water Balance in Wheat. In W. Day & R. K. Atkin (Eds.), *Wheat Growth and Modelling* (pp. 293–305). Springer.
- Ritchie, J. T. (1998). Soil water balance and plant water stress. In F. W. T. P. de Vries, G. Y. Tsuji, G. Hoogenboom, & P. K. Thornton (Eds.), *Systems Approaches for Sustainable Agricultural Development. Understanding Options for Agricultural Production* (pp. 41–54). Springer.
- Sadler, E. J., Gerwig, B. K., Evans, D. E., Busscher, W. J., & Bauer, P. J. (2000). Site-specific modeling of corn yield in the SE coastal plain. *Agricultural Systems*, 64(3), 189–207. [https://doi.org/10.1016/S0308-521X\(00\)00022-6](https://doi.org/10.1016/S0308-521X(00)00022-6).
- Saxton, K. E., Rawls, W. J., Romberger, J. S., & Papendick, R. I. (1986). Estimating generalized soil-water characteristics from texture. *Soil Science Society of America Journal*, 50(4), 1031–1036. <https://doi.org/10.2136/sssaj1986.03615995005000040039x>.
- Thorp, K. R., De Jonge, K. C., Kaleita, A. L., Batchelor, W. D., & Paz, J. O. (2008). Methodology for the use of DSSAT models for precision agriculture decision support. *Computers and Electronics in Agriculture*, 64(2), 276–285. <https://doi.org/10.1016/j.compag.2008.05.022>.
- USDA Natural Resources Conservation Service, Soil Survey Staff (2019). Soil Texture Calculator. Natural Resources Conservation Service, United States Department of Agriculture. <https://data.nal.usda.gov/dataset/soil-texture-calculator>. Accessed 2022-12-13.
- van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y. A., Padarian, J., Schaap, M. G., Tóth, B., Verhoef, A., Vanderborght, J., Ploeg, M. J., Weihermüller, L., Zacharias, S., Zhang, Y., & Vereecken, H. (2017). Pedotransfer functions in Earth System Science: Challenges and perspectives. *Reviews of Geophysics*, 55(4), 1199–1256. <https://doi.org/10.1002/2017RG000581>.
- Willmott, C. J. (1981). On the validation of models. *Physical Geography*, 2(2), 184–194. <https://doi.org/10.1080/02723646.1981.10642213>.
- Wösten, J., Pachepsky, Y., & Rawls, W. J. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251(3–4), 123–150. [https://doi.org/10.1016/S0022-1694\(01\)00464-4](https://doi.org/10.1016/S0022-1694(01)00464-4).

- Wu, R., Lawes, R., Oliver, Y., Fletcher, A., & Chen, C. (2019). How well do we need to estimate plant-available water capacity to simulate water-limited yield potential? *Agricultural Water Management*, *212*, 441–447. <https://doi.org/10.1016/j.agwat.2018.09.029>.
- Yang, J. M., Yang, J. Y., Liu, S., & Hoogenboom, G. (2014). An evaluation of the statistical methods for testing the performance of crop models with observed data. *Agricultural Systems*, *127*, 81–89. <https://doi.org/10.1016/j.agsy.2014.01.008>.
- Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture—a worldwide overview. *Computers and Electronics in Agriculture*, *36*(2–3), 113–132. [https://doi.org/10.1016/S0168-1699\(02\)00096-0](https://doi.org/10.1016/S0168-1699(02)00096-0).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.