Impact of calibrating a low-cost capacitance-based soil moisture sensor on AquaCrop model performance

Adla, Soham; Bruckmaier, Felix; Arias-Rodriguez, Leonardo F.; Tripathi, Shivam; Pande, Saket; Disse, Markus

DOI
10.1016/j.jenvman.2024.120248

Publication date
2024

Document Version
Final published version

Published in
Journal of Environmental Management

Citation (APA)
Adla, S., Bruckmaier, F., Arias-Rodriguez, L. F., Tripathi, S., Pande, S., & Disse, M. (2024). Impact of calibrating a low-cost capacitance-based soil moisture sensor on AquaCrop model performance. *Journal of Environmental Management*, *353*, Article 120248. https://doi.org/10.1016/j.jenvman.2024.120248

Journal of Environmental Management 353 (2024) 120248

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Impact of calibrating a low-cost capacitance-based soil moisture sensor on AquaCrop model performance

Soham Adla [a,1,*], Felix Bruckmaier [a], Leonardo F. Arias-Rodriguez [a], Shivam Tripathi [b], Saket Pande [c], Markus Disse [a]

[a] Chair of Hydrology and River Basin Management, Technical University of Munich, Arcisstrasse 21, 80333, Munich, Germany
[b] Department of Civil Engineering, Indian Institute of Technology Kanpur, Kanpur, 208016, Uttar Pradesh, India
[c] Department of Water Management, Delft University of Technology, 2628, CN Delft, the Netherlands

ARTICLE INFO

Handling editor: Jason Michael Evans

Keywords:
Low-cost soil moisture sensor
Calibration
AquaCrop
Crop modeling
Machine learning
Water productivity

ABSTRACT

Sensor data and agro-hydrological modeling have been combined to improve irrigation management. Crop water models simulating crop growth and production in response to the soil-water environment need to be parsimonious in terms of structure, inputs and parameters to be applied in data scarce regions. Irrigation management using soil moisture sensors requires them to be site-calibrated, low-cost, and maintainable. Therefore, there is a need for parsimonious crop modeling combined with low-cost soil moisture sensing without losing predictive capability.

This study calibrated the low-cost capacitance-based Spectrum Inc. SM100 soil moisture sensor using multiple least squares and machine learning models, with both laboratory and field data. The best calibration technique, field-based piece-wise linear regression (calibration $r^2 = 0.76$, RMSE = 3.13 %, validation $r^2 = 0.67$, RMSE = 4.57 %), was used to study the effect of sensor calibration on the performance of the FAO AquaCrop Open Source (AquaCrop-OS) model by calibrating its soil hydraulic parameters.

This approach was tested during the wheat cropping season in 2018, in Kanpur (India), in the Indo-Gangetic plains, resulting in some best practices regarding sensor calibration being recommended. The soil moisture sensor was calibrated best in field conditions against a secondary standard sensor (UGT GmbH. SMT100) taken as a reference ($r^2 = 0.67$, RMSE = 4.57 %), followed by laboratory calibration against gravimetric soil moisture using the dry-down ($r^2 = 0.66$, RMSE = 5.26 %) and wet-up curves respectively ($r^2 = 0.62$, RMSE = 6.29 %). Moreover, model overfitting with machine learning algorithms led to poor field validation performance. The soil moisture simulation of AquaCrop-OS improved significantly by incorporating raw reference sensor and calibrated low-cost sensor data. There were non-significant impacts on biomass simulation, but water productivity improved significantly. Notably, using raw low-cost sensor data to calibrate AquaCrop led to poorer performances than using the literature. Hence using literature values could save sensor costs without compromising model performance if sensor calibration was not possible. The results suggest the essentiality of calibrating low-cost soil moisture sensors for crop modeling calibration to improve crop water productivity.

1. Introduction

Surface irrigation, like other traditional irrigation methods, is inefficient due to deep percolation and non-uniform distribution of water (Pramanik et al., 2022). However, it may continue to remain the most extensively used irrigation method (Raine, 2006) due to its low-cost and energy requirements (Bjorneberg, 2013). Hence, technological improvements to improve irrigation management need to be relevant and correspondingly cost-effective, particularly when they are designed for smallholder farmers, who tend to be more economical in technology adoption (Singh et al., 2009).

Irrigation management techniques have combined either data from

* Corresponding author.
E-mail addresses: s.adla@tudelft.nl (S. Adla), felixbruckmaier@gmx.de (F. Bruckmaier), leonardo.arias@tum.de (L.F. Arias-Rodriguez), shiva@iitk.ac.in (S. Tripathi), s.pande@tudelft.nl (S. Pande), markus.disse@tum.de (M. Disse).
[1] Present address is Dept. of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Building 23, Stevinweg 1, 2628 CN Delft, The Netherlands.

https://doi.org/10.1016/j.jenvman.2024.120248
Received 13 November 2023; Received in revised form 8 January 2024; Accepted 27 January 2024
Available online 6 February 2024

satellites (Bastiaanssen et al., 2000) or ground sensor networks (Navarro-Hellín et al., 2015) with agro-hydrological modeling (Chiara and Marco, 2022). Precise irrigation water management requires accurate measurements as well as a comprehensive biophysical process understanding of crop response to water at different crop growth stages (Kisekka et al., 2022). Crop water simulation models describe the crop growth and production responses to the soil-water environment (C. Zhang et al., 2022). Such models are often data intensive, requiring numerous input variables and parameter values, which may not be available for different crops and environments (Vanuytrecht et al., 2014). Similarly, though commercial sensors, processors and communication components may be available, they need to be cost-effective, easy to access and maintainable for effective agricultural applications (Pramanik et al., 2022; Rodríguez-Robles et al., 2020).

Soil moisture (or volumetric water content, VWC) is a key variable involved in precision agriculture and agricultural sustainability (Kisekka et al., 2022), which can be an input to crop growth models. Soil moisture sensing technologies can have low-utilization due to farmers' economic limitations and reluctance towards such investment (Srbinovska et al., 2015). Further, soil moisture sensors need to be calibrated at each site because of the effects of the variability in soil properties on the sensor outputs (Peddinti et al., 2020). Researchers have highlighted the importance of field calibration as undisturbed soil samples are representative of field conditions where the sensors might be used (Robinson et al., 2003). However, soil moisture sensing is sensitive to environmental factors like temperature, salinity, bulk density, organic matter and clay content (Kargas and Soulis, 2012; Matula et al., 2016). Hence, manufacturers generally perform indoor calibration with sieved, uniformly packed soils at regulated VWCs and temperature conditions, particularly in homogeneous coarse soils like some sands and loams (Feng and Sui, 2020).

Studies have used irrigation scheduling with calibrated low-cost capacitance sensors to detect crop water stress (Thompson et al., 2007a) and irrigation thresholds computed using measured VWC-based indices(Thompson et al., 2007b). Other studies have used expensive sensors in combination with crop modeling (Lu et al., 2021). The challenge is to combine accurate (calibrated) soil moisture sensing at low costs and parsimonious crop models without losing predictive power (Landau et al., 2000). This paper addresses this using four research questions. RQ1 pertains to LC soil moisture sensor performance on the field, and RQs 2 to 4 pertain to using LC soil moisture data for parsimonious crop model calibration:

1. Can low-cost (LC) soil moisture sensors be calibrated using more sophisticated techniques such as machine learning algorithms to improve their field performance?
2. Do raw data from LC soil moisture sensors have any advantage over literature values for parsimonious crop model calibration?
3. What is the difference in performance when such a model is calibrated using raw data from a LC vis-à-vis a reference soil moisture sensor?
4. What is the effect of calibrating the LC soil moisture data against a reference sensor on crop model performance?

The novelty of this study lies in combining data from LC soil moisture sensors with parsimonious crop modeling to evaluate crop model performance in simulating canopy cover, VWC and water productivity (WP) using a two-step calibration process. Firstly, low-cost (capacitance based) soil moisture sensors were calibrated against uncalibrated reference (TDR-FDR) sensors via multiple models (least squares and machine learning algorithms). Consequently, these calibrated VWC data were used for calibrating the soil hydraulic parameters of a parsimonious crop model to assess improvements in WP. Such a laboratory-field-simulation study based on a food critical region in the Majority World, which combines low-cost soil moisture sensor calibration (using multiple models) with parsimonious crop modeling, has not been undertaken yet, to the best of the authors' knowledge.

2. Material and methods

2.1. Study area and wheat cropping season

2.1.1. Experimental site

The study was conducted during the wheat cropping season of 2018 in Kanpur, which is representative of an intensively managed rural landscape in the Indo-Gangetic plains (Gupta et al., 2019). Measurements were collected in a 20 m × 30 m experimental field (Fig. 1) at the Indian Institute of Technology Kanpur, Kanpur, India (26°30′56.8″N, 80°13′47.3″E and altitude of about 126 m above mean sea level). The experimental site falls within the sub-tropical climate zone with an average annual rainfall of 833.5 mm, 92.5 % of which falls within the monsoon season (June to September); the other two seasons are the cold season from November–February and the hot season from March–June (Sankararamakrishnan et al., 2008).

2.1.2. Description of the wheat cropping season

The Indian spring wheat variety K7903 (improved *Halna*) (Dwivedi et al., 2019; Kumar et al., 2012) was sown in check basins (or 'plots', each 3 m × 3 m) on January 5, 2018. The seeds were sown with a row-spacing, depth and planting density of 10 cm, 5 cm and 363 plants/sq.m respectively. *Halna* is a very late sown, short duration, drought tolerant variety of wheat (Dwivedi et al., 2019; Kumar et al., 2017).

Agricultural management was performed based on local farming practices, that included fertilizer (urea) application during tillering (27 days after sowing, or DAS) and 4 irrigation applications (25 DAS, 45 DAS, 59 DAS and 74 DAS, given in Table A1) totaling 220.2 mm, to supplement the 50.4 mm rainfall during the cropping season. There was a flooding event which occurred just following the final irrigation, totaling 262.5 mm of water. Hence, the total water input was 533.1 mm. The crop was harvested on April 15, 2018 (101 DAS).

2.2. Soil moisture sensing methods

2.2.1. Selecting appropriate soil moisture sensing techniques

The various techniques for soil moisture sensing include neutron thermalization (Chanasyk and Naeth, 1996), water content reflectometer (Chandler et al., 2004; Kargas and Soulis, 2019), time domain reflectometry (TDR) (Robinson et al., 2003; Topp et al., 1980), time domain transition (TDT) (Blonquist et al., 2005), electrical impedance (Cosh et al., 2005; Gaskin and Miller, 1996), frequency domain reflectometry (FDR) (Ojo et al., 2015), electrical capacitance (Zotarelli et al., 2011), and sensing using electrical resistance blocks (Cummings and Chandler Jr, 1941) or tensiometers (Muñoz-Carpena et al., 2005).

Capacitance-type soil moisture sensors are more widely used for decision support systems in irrigated agriculture (Fares and Alva, 2000; Gallardo et al., 2020) due to their lower cost, robustness, precision, and low power and maintenance requirements (Jones et al., 2005; Rosenbaum et al., 2011; Spelman et al., 2013; Visconti et al., 2014). However, despite performing well in laboratory conditions, they may exhibit sensor-to-sensor variability in field conditions (Bogena et al., 2017; Rosenbaum et al., 2010; Spelman et al., 2013), and consequently require site-specific calibration to be able to provide reliable VWC measurements (Kisekka et al., 2022; Peddinti et al., 2020). The.

The WaterScout SM100 soil moisture sensor (Spectrum Technologies, Inc., Plainfield, IL, USA) was chosen as the capacitance based 'low-cost (LC)' soil moisture sensor (Spectrum Technologies, 2014) investigated in the study. The accuracy and operating range of SM100 are 3 % (for EC $<$ 800 mS m^{-1}) and 0.5 °C–80 °C respectively (Spectrum Technologies, 2014). Laboratory calibration (with repacked sands and silty loams) has improved its VWC accuracy (RMSE, Raes et al., 2018) to between 1.63 % and 2.97 % (Adla et al., 2020). The manufacturer's

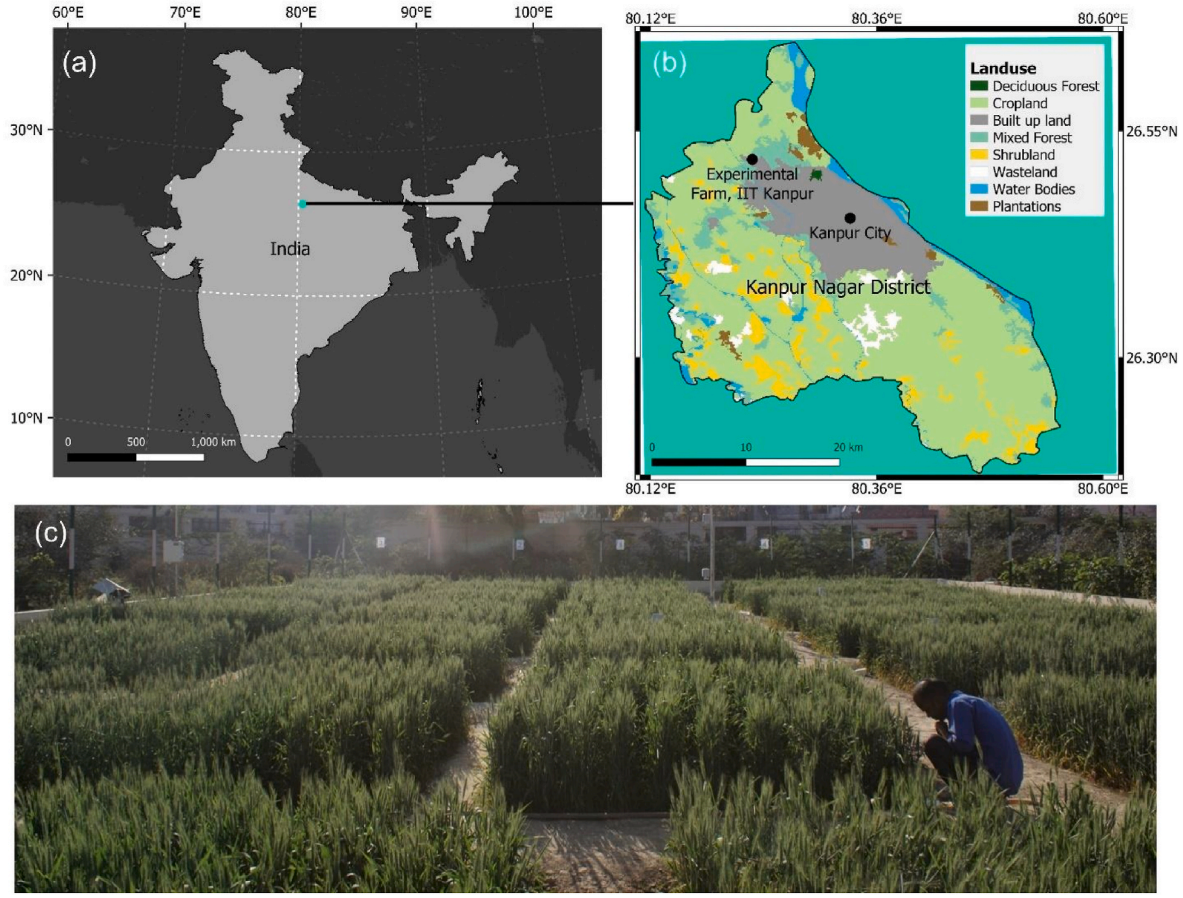

Fig. 1. (a): Location of Kanpur in India, (b) Location of the Indian Institute of Technology (IIT) Kanpur in the Kanpur district (with a landuse map), (c) Experimental wheat farm in IIT Kanpur, with check-basins (plots) irrigated using flood irrigation. Six plots with soil moisture monitoring using both Ref. and LC sensors were chosen for the study.

calibration equation (D. Kieffer, personal communication, September 5, 2018) is given below:

$$VWC\ (\%) = -19.13 + 0.016\ (V_R) \text{ if } V_R \leq 0.34 \quad (1a)$$

$$VWC\ (\%) = 263.52 - 0.385\ (V_R) + 1.14 \times 10^{-4}\ (V_R^2) \text{ if } V_R > 0.34 \quad (1b)$$

where VWC is measured in %, and V_R is the voltage ratio of the output and input voltages (V_{out}/V_{in}).

The SMT100 sensor (manufactured by Umwelt-Geräte-Technik GmbH, Müncheberg, Germany) was chosen as the secondary standard soil moisture sensor (Nakra and Chaudhry, 2006) to calibrate the LC SM100 sensor (see Section 2.4.1.2.). It is henceforth called the 'reference (Ref)' sensor due to its superior sensing technology which combines the higher accuracy of a TDR system with the cost effectiveness of FDR sensing with an oscillating frequency of 340 Hz in air and 150 MHz in water (Bogena et al., 2017; Umwelt-Geräte-Technik GmbH, 2017). Its accuracy for VWC and temperature are 3 % and ±0.2 °C respectively, and it operates within the temperature range of −40 °C–80 °C (Umwelt-Geräte-Technik GmbH, 2017). Sensor specific laboratory calibration with (non-soil) materials of known apparent dielectric permittivity can improve the VWC accuracy (RMSE) to range between 0.21 % and 1.30 % (Bogena et al., 2017). The manufacturer's calibration equation is identical to the one proposed by (Topp et al., 1980):

$$VWC\ (\%) = -5.30 \times 10^{-2} + 2.92 \times 10^{-2}\epsilon_r - 5.50 \times 10^{-4}\epsilon_r^2 + 4.30 \times 10^{-6}\epsilon_r^3 \quad (2)$$

where ε_r is the sensed dielectric constant.

Since the only common output across both SM100 and SMT100 sensors was VWC, the uncalibrated SM100 VWC values were calibrated on the SMT100 sensor's output VWC values. Details of sensor outputs are mentioned in the Appendix.

A short description of the sensors (along with their costs) is given in Table 1, with more details available in the Appendix.

2.2.2. *Selecting calibration techniques for the LC capacitance sensor*

2.2.2.1. Least squares estimate based regression. Studies have calibrated capacitance-based soil moisture sensors to improve performance, both in laboratories (Adla et al., 2020; Bello et al., 2019; Nagahage et al., 2019; Placidi et al., 2021) and in the field (Rudnick et al., 2015; Singh et al., 2018). Most of the calibration equations reported in the literature for low-cost sensors are least-squares estimates (e.g., linear, logistic, hyperbolic, logarithmic, exponential, or polynomial). Based on a visual inspection of the data, piecewise linear regression, power law regression

Table 1
Details of soil moisture sensors used in the study. Costs are derived from quotations received by the authors.

Soil moisture sensor (manufacturer)	Measurement technique	Price (quotation)	Nomenclature in the study
SM100 soil moisture sensor (Spectrum Technologies, Inc.)	Capacitance based	$89	Low-cost
SMT100 soil moisture sensor (UGT GmbH)	TDR-FDR based	$155	Reference

and polynomial regressions (degree 2 and 3) were used to calibrate the LC capacitance based SM100 sensor.

2.2.2.2. Machine learning based regression. Diverse machine learning algorithms were also applied, including standard algorithms for regression applications: linear regression (LR), support vector regression (SVR), random forest regression (RFR) and multilayer perceptron neural networks (MLP). LR models have been applied extensively in soil moisture applications (Qiu et al., 2003; Teng et al., 1993); and are still relevant due to their simplicity and easy interpretation (García et al., 2016; Lee et al., 2019). The Support Vector Machine (SVM) algorithm defines optimal hyperplanes in a high or infinite dimensional space which can be used for classification or regression (Vapnik et al., 1996). SVR has been applied in soil moisture related studies (Gill et al., 2006; Liu et al., 2010; Yu et al., 2012). The RFR algorithm (Breiman, 2001) is based on averaging non-correlated decision trees for variance reduction and avoidance of overfitting. Its simplicity in training and tuning has made it popular in current regression applications of soil moisture (Carranza et al., 2021; Qingling et al., 2019; Srivastava et al., 2021; H. Zhang et al., 2022). However, calibration hyperparameters can be computationally expensive due to its many possibilities. The MLP is an algorithm based on the typical architecture of a neural network, hence it is a nonlinear statistical model which has unknown parameters (weights) meant to be tuned to make the model fit well to training data using back-propagation equations in multiple hidden layers. It has also been applied in soil moisture applications extensively (Chai et al., 2009; Gu et al., 2021; Yan et al., 2010; Yu et al., 2012, p. 20). All the algorithms were applied using the Scikit-learn tool library (v.10.2) (Pedregosa et al., 2011) in Python (v.3.8.5). Hyperparameters of each algorithm were calibrated with default values available in Scikit-learn and evaluated with a 10-fold cross validation (CV, k = 10), and are listed in Table A3. This process avoided skewing results in validation due to random sampling in the training process. The controlling metrics were the coefficient of determination (R^2, Raes et al., 2018) and the root mean squared error (RMSE). Hastie et al. (2009) and Arias-Rodriguez et al. (2021) provide detailed definitions and an application of all these machine learning models, respectively.

2.3. Parsimonious crop modeling

2.3.1. Selecting a parsimonious crop model

Crop growth models can simulate physiological processes (Chenu et al., 2009; Yin et al., 2003), and crop growth behavior in the field (Keating et al., 2003; Robinson et al., 2003; Steduto et al., 2009). Consequently, they can inform management decisions regarding water and nutrients and explore the feasibility of new cropping systems (Asseng et al., 2014; Silva et al., 2017). The more popular models used for simulating crop growth and yield production of wheat under different soil water conditions and irrigation scenarios are: APSIM (Ahmed et al., 2016; Chen et al., 2010), DSSAT-CERES-Wheat (Attia et al., 2016; Zhou et al., 2018), FAO AquaCrop (Iqbal et al., 2014; Toumi et al., 2016), RZWQM2 (Saseendran et al., 2015; Zheng et al., 2020), and SWAP (Eitzinger et al., 2004; Wang et al., 2021).

Crop modeling environments like DSSAT and APSIM can require 211 and 292 parameter inputs respectively (Soltani and Sinclair, 2015). This increased complexity enables such models to address complex research questions (Todorovic et al., 2009), but also amplifies errors and uncertainties (Silva and Giller, 2020). Further, data intensive models with numerous inputs and parameter requirements may not be practical in data-scarce regions (Vanuytrecht et al., 2014; Varella et al., 2010) like the Majority World (Graves et al., 2002; Jones et al., 2012).

The water-driven FAO AquaCrop model has been envisioned to balance simplicity, accuracy and robustness (Vanuytrecht et al., 2014). Consequently, it relies on considerably fewer and relatively easier to measure inputs (about 19, according to FAO, 2016; Raes, 2017). Nevertheless, FAO AquaCrop's performance is comparable to data-intensive models (Babel et al., 2019; Quintero and Díaz, 2020; Todorovic et al., 2009). Moderate to good simulation results have also been reported for wheat and regions of water scarcity (Andarzian et al., 2011; Huang et al., 2022; Kale et al., 2018; Singh et al., 2013; Zhang et al., 2013).

2.3.2. Description of the crop-water model FAO AquaCrop

FAO AquaCrop simulates daily yield loss in response to soil water depletion in the root zone, based on the approach of Doorenbos and Kassam (1979). Four main variables are calculated consecutively and on a daily basis through individual equations; i.e., accumulated canopy cover (CC), daily plant transpiration, accumulated aboveground biomass, and accumulated final dry yield. These model variables are mainly interconnected by empirical factors, such as the water productivity (WP) and harvest index (HI), which convert transpiration to biomass and biomass to yield, respectively. Water productivity only incorporates plant transpiration to account for the confounding effect of the nonproductive consumptive water use (Steduto et al., 2009). Moreover, WP is normalized for atmospheric CO_2 concentration and climate (Raes et al., 2018), and hence makes the model more robust and generalizable (Steduto et al., 2009).

Model set-up requires user input for climate parameters, 'non-conservative' (i.e., spatio-temporally variable) model parameters, and, where applicable, an irrigation schedule and the groundwater water table. The model also needs daily data on precipitation, reference evapotranspiration, and minimum and maximum temperature. Non-conservative model parameters relate to crop phenology, soil conditions, and field management practices (Raes et al., 2018; Steduto et al., 2012). This study used the MATLAB-based AquaCrop OpenSource (AquaCrop-OS) v.6.1 tool (Foster et al., 2017).

2.4. Measurement methodology

An automatic weather station at the experimental field measured precipitation, barometric pressure, global solar radiation, wind speed and direction. It also recorded air temperature and relative humidity at two different heights (2 m and 3 m above the ground). These data were measured at 15-min intervals and aggregated to daily values. During the cropping season, the average daily temperature and relative humidity (both at 2 m) ranged from 9.3 °C to 30.7 °C and from 35.3 % to 93.9 %, respectively.

Soil texture classification was conducted by the UGT Sedimat 4–12 instrument, which determines the particle size distribution in mineral soils based on the Köhn method (König et al., 2005).

The observed crop growth parameters included phenological parameters like days to emergence, start of flowering, start of senescence and maturity, length of flowering (days). The observed soil related parameter was maximum rooting depth. The observed crop management parameter was plant population density.

The observations used for calibrating model parameters were leaf area index, VWC and above ground biomass. Leaf Area Index (LAI) was measured eight times during the season (36, 44, 55, 59, 68, 83, 90 and 101 DAS) using the LAI-2200C plant canopy analyzer manufactured by LI-COR Biosciences. It was converted into Canopy Cover values using an empirical equation for wheat (Nielsen et al., 2012). Surface soil moisture was determined at the center of each plot at 15-min intervals by Ref. SMT100 and LC SM100 sensors installed at a depth of 5 cm below the soil surface (Section 2.2.1). Both datasets were aggregated to the daily time-step. Above ground dry biomass was measured using crop cutting experiments (Singh, 2014) from a representative 1 sq. m section for each plot. Additionally, water productivity (WP_{ET}), introduced in Section 2.5.2, was calculated using actual evapotranspiration (ET_a) measured using microlysimeters (Kumar, 2019).

Six plots were used for all the analyses, as they had both Ref. and LC soil moisture data, as well as the observations required for crop

modeling. VWC measurements for both the sensors were available on average for 61.2 days (s = 3.0 days) out of the 100-day cropping season across the six plots.

2.5. Calibration methodology of low-cost soil moisture sensors

2.5.1. Strategy to calibrate-validate soil moisture data

Capacitance based soil-moisture sensors have been calibrated both with repacked (Adla et al., 2020; Nagahage et al., 2019; Placidi et al., 2020) or undisturbed (Bello et al., 2019) soil samples inside the laboratory, as well as in the field (Rudnick et al., 2015; Singh et al., 2018). In the laboratory, calibration of soil moisture sensors is generally carried out either using the substantially faster 'wet-up' or wetting (downward or upward infiltration, taking <1 day) or slower 'dry-down' or drying (taking a few weeks) processes (Burns et al., 2014).

Fig. 2 describes the overall study workflow, contextualizing each research question. The sensors used are displayed at the top-left corner of the figure. Throughout the study, the data corresponding to the LC SM100 and Ref. SMT100 sensors are represented in blue and green colors respectively. The simulations corresponding to the calibrated LC soil moisture sensor data are represented in a deeper blue compared to the lighter blue representing raw LC sensor data. The workflow related to the LC soil moisture sensor (SM100) calibration is illustrated in the inset "Soil moisture sensor calibration", in Fig. 2. To address RQ1, the LC sensor data were calibrated against either a primary standard (gravimetric water content, Section 2.4.1.1.) or a secondary standard (the Ref. SMT100 data, Section 2.4.1.2.) - this is mentioned in the respective bubbles, "Lab Cal." (laboratory calibration) and "Field Cal." (field calibration). The number of paired data points (sensor and primary/secondary standard) used in both calibration approaches were n = 100 and n = 65 for laboratory and field calibration, respectively. The following process resulted in selecting the "best" LC calibration model to be used for crop modeling (Section 2.5.2.).

2.5.1.1. Laboratory calibration of LC SM100 sensor. A previous study calibrating the LC SM100 sensor (Adla et al., 2020) used piece-wise linear regression functions (PWLFs) (Jekel, 2017, p. 20) against gravimetric VWC (primary standard), for four soils (number of data points, n = 400). The study calibrated five sensors in controlled laboratory conditions with the wet-up curve (similar to (Matula et al., 2016)) to account for sensor-to-sensor-variability and improve cost-effectiveness in terms of saved time and energy resources (Burns et al., 2014). This study used a subset (n = 100) of the data from Adla et al. (2020) which represented the soil sampled from the current study site. LC sensor values were calibrated against gravimetric VWC using the different algorithms described in Section 2.2.2. R^2 and RMSE were used to quantify the

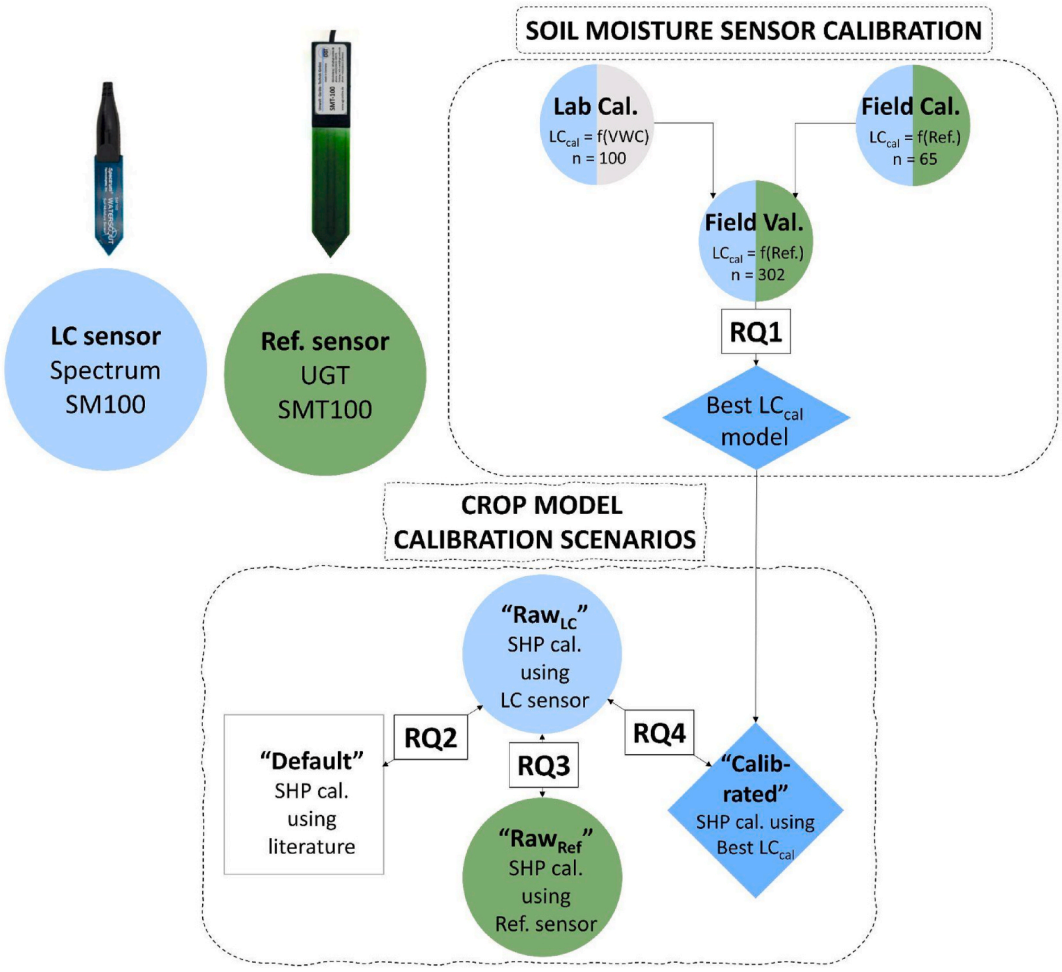

Fig. 2. Overview of the study methodology, contextualizing the four research questions (RQs). Sensor details are given at the top-left corner. The "Soil moisture sensor calibration" of the LC SM100 sensor was conducted in the laboratory ("Lab Cal.") using gravimetrically determined VWC, and in the field ("Field Cal.") using the Ref. SMT100 sensor. All the developed models were then validated in the field ("Field Val.") to select the most suitable model, "Best LC_{cal} model". The "Crop model calibration" either used the literature ("Default"), raw VWC values from both Ref. and LC sensors ("Raw_{Ref}" and "Raw_{LC}"), or the Best LC_{cal} model ("Calibrated") to calibrate the soil hydraulic parameters (SHPs) of the AquaCrop model. The Best LC_{cal} rhombus in "Soil moisture sensor calibration" is a flowchart decision, and the "Calibrated" diamond in "Crop model calibration" is a representation followed in Fig. 5.

calibration performance in all cases.

2.5.1.2. Field calibration of LC SM100 sensor. Out of the six plots used for the study, the plot with the highest (significant) linear correlation between the aggregated daily values of Ref. SMT100 and LC SM100 sensor (n = 65, R^2 = 0.76) was selected for calibration. Due to the impracticality of frequent gravimetric measurements and its superior sensing technology, the Ref. SMT100 sensor was selected as a secondary standard (Section 2.2.1.). Consequently, the daily aggregated LC SM100 data were calibrated using the daily aggregated Ref. SMT100 data, using the algorithms listed in Section 2.2.2. Calibration performance in each case was quantified by R^2 and RMSE indices.

2.5.1.3. Field validation of laboratory and field calibrations. Each calibration algorithm developed using both laboratory and field data were validated on independent data from the five remaining plots (n = 302). The Ref. SMT100 data measurements were taken as the secondary standard used for validating the field data. R^2 and RMSE were used as the performance indicators for the validation process.

2.6. Crop model calibration scenarios

2.6.1. AquaCrop calibration method

AquaCrop guidelines recommend calibrating all non-conservative and non-observed model parameters sequentially, with appropriate objective functions (AquaCrop. Tutorials., 2016; Raes et al., 2018; Steduto et al., 2012). Accordingly, model simulation is sequentially improved using intermediately computed model variables like canopy cover (CC), volumetric water content (VWC) and biomass (BM) by adjusting different sets of model parameters, until the simulation of the harvested yield agrees, visually and statistically, with observations within an acceptable error range (Raes et al., 2018).

In this study, eight (non-conservative) crop growth parameters of the AquaCrop-OS model were calibrated: phenological parameters (initial canopy size of the emergent seedling CC_o, maximum canopy cover CC_x, and canopy growth coefficient CGC) and soil hydraulic parameters (readily available water REW, saturated soil hydraulic conductivity K_{sat}, VWC at permanent wilting point θ_{PWP}, field capacity θ_{FC} and saturation θ_s).

2.6.2. Different calibration scenarios used in the study (default, raw and calibrated)

The inset "Crop model calibration" in Fig. 2 illustrates the FAO AquaCrop modeling scenarios used to address RQs 2 to 4. Crop growth parameters (CC_o, CC_x, CGC) were calibrated using a trial-and-error method (Liang et al., 2017; Ma et al., 2020) to maximize the Pearson correlation coefficient (r) and minimize the RMSE between the measured and simulated CC values for each plot (Raes et al., 2018). The soil hydraulic parameters (SHPs), θ_{WP} (%), θ_{FC} (%), θ_s (%), K_{sat} (mm/day), and REW (mm), were also calibrated to respectively maximize and minimize r and RMSE between simulated and observed (Ref. SMT100) soil moisture data. Soil Hydraulic Parameter (SHP) calibration was conducted under the following scenarios:

- **The "Default" scenario:** SHPs were calibrated based on average values from the literature (Rawls and Brakensiek, 1989; Gupta et al., 2021), corresponding to the classified soil texture (silty-loam).
- **The "Raw" scenarios (for both Ref. and LC sensor):** SHPs were calibrated with the raw (uncalibrated) LC and Ref. soil moisture sensor data using a trial-and-error method (Liang et al., 2017; Ma et al., 2020).
- **The "Calibrated" scenario (only for LC sensor):** SHPs were calibrated using the "best" LC_{cal} model.

The difference in AquaCrop model performance between the Default and Raw_{LC} scenarios would illustrate the advantage of using raw LC data over literature values (RQ2). The difference between the performance of the models calibrated in the Raw_{LC} and Raw_{Ref} scenarios would address RQ3, indicating the relative loss in crop model performance when LC sensor data is used for calibration vis-à-vis reference soil moisture sensor data. The difference between the "Calibrated" (Cal_{LC}) and "Raw" (Raw_{LC}) model performances would quantify the additional impact (on crop model performance) of using calibrated vis-à-vis raw LC soil moisture data (RQ4).

Above ground dry biomass (BM) simulated by AquaCrop was compared with the observed biomass, using the mean absolute error (MAE, Witten et al., 2011), RMSE, NRMSE (normalized RMSE, Raes et al., 2018), and percentage bias (PBIAS, Sorooshian et al., 1993) indicators. Additionally, Water Productivity (WP) was simulated for each scenario to understand the "efficacy of the crop production processes in relation to their required water consumption" (Van Halsema and Vincent, 2012). This study used ET Water Productivity or WP_{ET}, defined as the following (Van Halsema and Vincent, 2012):

$$WP_{ET} = Yield \ / \ ET_a \ (\mathrm{kg}/\mathrm{m}^3 \text{ or } \mathrm{kg}/\mathrm{kg}) \quad (3)$$

where *Yield* is the final (simulated or observed) crop yield (kg/m^2) and ET_a (m) is the cumulative actual evapotranspiration during the cropping season. The simulated WP_{ET} was computed by extracting the yield and seasonal evapotranspiration values simulated by AquaCrop. Observed WP_{ET} was computed using the observed biomass, the default harvest index for wheat, and ET_a estimated on the field using microlysimeters (Kumar, 2019).

3. Results and discussion

3.1. Description of the cropping season

Fig. 3 illustrates the seasonal evolution of the daily water demand (depicted by the FAO-56 reference ET_o, from (Allen et al., 1998)), water supply (depicted by rainfall and applied irrigation during the season), and soil moisture (VWC) monitoring by both low-cost (LC) SM100 and reference (Ref.) SMT100 soil moisture sensors. The soil texture, with 15.4 % sand and 66.3 % silt, was determined to have silty-loam texture (USDA classification). Moreover, the soil was relatively homogenous in depth and the groundwater table was deep enough to prevent capillary rise from influencing soil moisture measurements.

3.2. Calibration of low-cost soil moisture sensors

Table 2 details the results of the LC sensor calibration and validation procedure (Section 2.4). The sub-table on the left and right (with 6 columns each) describe the calibration-validation performance in the laboratory and field respectively, using R^2 and RMSE as the performance indicators. The validation in both cases is done on the same independent field data (n = 302). Fig. 4 plots all the calibration curves preserving the index from Table 2, highlighting the selected "best" LC_{cal} model.

During both laboratory and field calibrations, the performance of the machine learning algorithm Random Forest Regression (RFR) was quantitatively the best among all the algorithms included in the study (R^2 = 0.98 and RMSE = 1.26 % for laboratory calibration; R^2 = 0.93 and RMSE = 1.67 % for field calibration). However, overfitting was observed (index 6 in Fig. 4), likely due to the limited tuning of the hyperparameters due to limited computational resources. Though RFR avoids overfitting from original decision trees (Hastie et al., 2009), generalization error variance decreases as more trees are added to the algorithm, without a change in the generalization bias. Though this study used the default *n_estimators* = 100 (Pedregosa et al., 2011), different options of decision trees should be examined, particularly searching for lower *n_estimators* using a GridSearch.

As a consequence, the RFR-laboratory model had inadequate

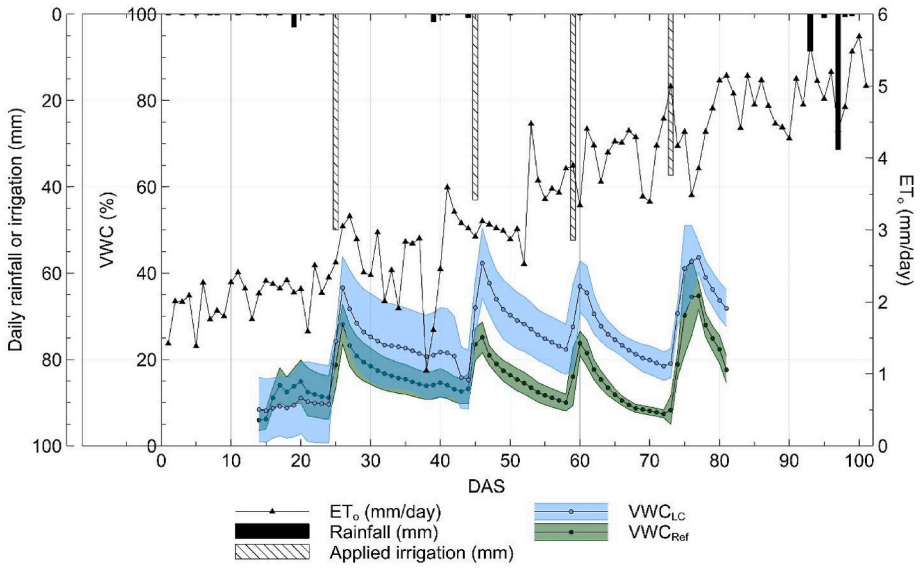

Fig. 3. Variation of daily values of FAO-56 reference evapotranspiration (ET_o), input water through rainfall or irrigation application, and soil moisture measurements (both reference SMT100 and low-cost SM100 sensors) during the wheat cropping season. Soil moisture averages over the field are depicted as bubbles, with their standard deviations as ranges.

Table 2
Overview of the performance indices for calibrating the LC sensor in the laboratory (left) and in the field (right). The different calibration techniques are 0. Piecewise linear function (PWLF), 1. polynomial of degree 2, 2. polynomial of degree 3, 3. Power law, 4. Linear regression (LR), 5. Support Vector Regression (SVR), 6. Random Forest Regression (RFR), 7. Multi-Layer Perceptron regression (MLP). Laboratory and field models 0. to 7. use the wet-up and dry-down curves respectively to calibrate the LC sensor. Models 8. and 9. correspond to laboratory calibrations from independent research with the dry down curve. Gedilu (2020) calibrated the LC sensor using the same field's soil, and Rai (2012) calibrated a nearby silty-loam soil, both using dry-down curves. The Best LC_{cal} model was the PWLF-field model (number 0.), highlighted in bold. n is the number of data points, R^2 is the coefficient of determination calculated during model development and RMSE is the root mean squared error.

	Calibration models	Calibration in the laboratory, and validation in the field						Calibration in the field and validation in the field					
		Laboratory Calibration (against gravimetric VWC)			Field Validation (against Ref. sensor)			Field Calibration (against Ref. sensor)			Field Validation (against Ref. sensor)		
		n	R^2	RMSE (%)	n	R^2	RMSE (%)	n	R^2	RMSE (%)	n	R^2	RMSE (%)
	Manufacturer's calibration	100	0.66	7.58	302	0.62	11.99	65	0.76	11.49	302	0.62	11.99
0	Best LC_{cal}: PWLF	100	0.95	1.96	302	0.42	7.98	**65**	**0.76**	**3.13**	**302**	**0.67**	**4.57**
1	Polynomial (d = 2)	100	0.88	2.92	302	0.00	11.28	65	0.76	3.18	302	0.64	5.09
2	Polynomial (d = 3)	100	0.93	2.16	302	0.12	9.69	65	0.76	3.18	302	0.66	4.66
3	Power Law	100	0.81	3.67	302	0.54	7.31	65	0.76	3.18	302	0.65	4.97
4	LR	100	0.66	4.93	302	0.62	6.29	65	0.76	3.19	302	0.62	5.43
5	SVR	100	0.87	3.09	302	0.10	9.12	65	0.62	3.97	302	0.47	5.84
6	RFR	100	0.98	1.26	302	0.20	8.88	65	0.93	1.67	302	0.62	4.87
7	MLP	100	0.50	5.97	302	0.62	6.85	65	0.71	3.50	302	0.62	4.70
8	Gedilu (2020)	Independent studies with dry-down			302	0.61	6.06						
9	Rai (2012)	curve			302	0.66	5.26						

validation performance (R^2 = 0.20, RMSE = 8.88 %). Moreover, the RFR-field and RFR-laboratory models were respectively non-monotonic and monotonically non-decreasing functions, implying that the calibrated LC VWC values could also decrease (or not increase, respectively) with increasing Ref. VWC. Both the model behaviors were not realistic. Consequently, the RFR algorithm was excluded from further analyses.

The piece-wise linear regression function (PWLF) had the next best performance during calibration: laboratory R^2 = 0.95, RMSE = 1.96 %; and field R^2 = 0.76, RMSE = 3.13 %. Further, during both the respective validation cases, the PWLF performed adequately well when compared to the relative performances of the RFR models (PWLF-laboratory: R^2 = 0.42, RMSE = 7.98 %, PWLF-field: R^2 = 0.67, RMSE = 4.57 %). This piece-wise (segmented) behavior is also a feature of the manufacturer's calibration equation which has linear and quadratic segments (Kieffer, personal communication, September 5, 2018).

Fig. 5 shows scatter plots between measured/calibrated LC SM100 values and the measured Ref. SMT100 values for (a) laboratory calibration, (b) validation of the laboratory-calibrated PWLF model, (c) field calibration, and (d) validation of the field-calibrated PWLF model. Raw and calibrated data are represented by cyan bubbles and deep blue diamonds respectively.

Some best practices for soil moisture sensor calibration were derived from this comparative analysis of calibration models, which can be more specific to capacitance-based sensors used in silty-loam agricultural soils:

- **Field calibration of fewer sensors may be more robust than laboratory calibration of multiple sensors:** Field calibrations with undisturbed soils may be more robust for field applications. The field validation performance (both R^2 and RMSE) of laboratory-calibrated models was significantly poorer than field-calibrated models, supporting previous recommendations of on-site calibration using undisturbed soil samples (Feng and Sui, 2020).

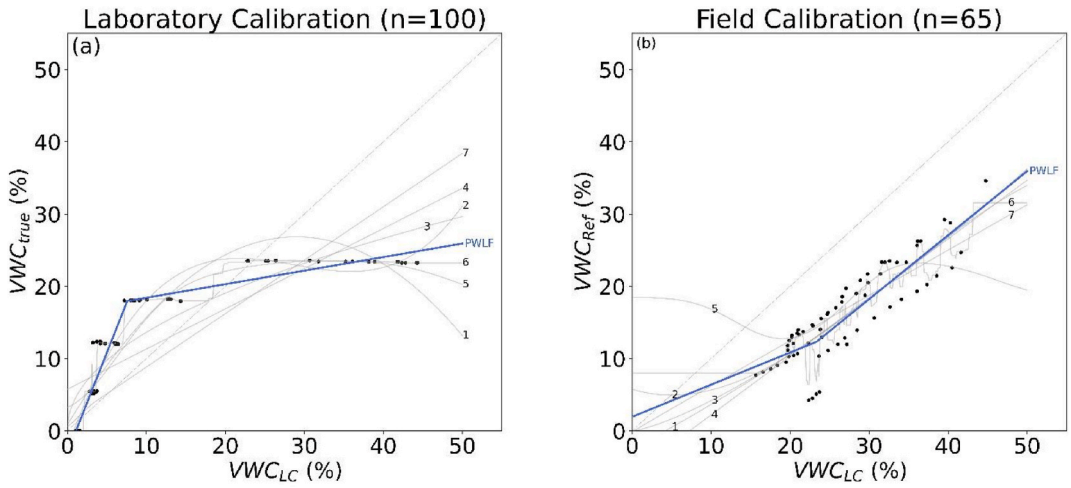

Fig. 4. Comparison of the different calibration equations developed using (a) lab and (b) field data. The index numbers correspond to Table 2. In both cases, the data used for calibration are given as points, and the "Best" model (in blue) is PWLF.

- **If field calibration is not possible, calibration should be done using the dry-down curve:**

Calibrating LC sensors in the laboratory using the dry-down curve rather than the wet-up curve leads to superior performance for field applications. The laboratory validation results corroborated with previous literature which claimed superior laboratory calibration accuracy using dry-down curves, particularly for finer textured soils (Burns et al., 2014). An earlier study which calibrated 3 LC SM100 sensors with a similar soil using the dry-down curve in the laboratory was found to have the best validation performance (R^2 = 0.66, RMSE = 5.26 %) among all laboratory calibrated models (Rai, 2012). Similarly, a linear dry-down calibration equation developed in the laboratory by Gedilu (2020) using only one LC SM100 sensor using repacked soil from the experimental field, also performed well (R^2 = 0.61 and RMSE = 6.06 %). These models performed better than all the laboratory models in this study (models 0 to 7), which were developed using the wet-up curve with five LC sensors. This suggests that capturing dry-down dynamics can compensate for the sensor-to-sensor variability even for repacked soils inside the laboratory. Further, the dry-down curve is more representative of the soil water dynamics during the dry non-monsoonal *rabi* cropping season in northern India, where rainfall majorly occurs during the monsoon *kharif* season (Sankararamakrishnan et al., 2008).

- **Overfitting during calibration can lead to loss in robust application:**

Overfitting was observed in both laboratory and field RFR models, both of which performed poorly during validation. Machine learning models, applied to non-linear hydrological processes have had limitations of overfitting and reasonable justification of results (Elshorbagy et al., 2010). The data in this study suggest that monotonic (preferably "gently sloped") curves are more robust for soil moisture calibration. Overfitting using ML techniques is possibly also due to the relatively lower number of samples that exemplify soil moisture sensor calibration. This supports previous studies which mainly report linear and polynomial calibration curves (Bello et al., 2019; Deng et al., 2020; Gedilu, 2020; Nagahage et al., 2019; Placidi et al., 2020; Rudnick et al., 2015; Singh et al., 2018; Thompson et al., 2007b, 2007a).

With respect to research question 1 (RQ1), the PWLF model calibrated using field data was chosen as the "best" LC_{cal} model due to the flexibility of multiple segments without compromising the positive correlation expected between data from two electromagnetic sensors in the same soil. This suggests that simpler models such as PWLF are more robust in calibrating LC soil moisture sensors than more sophisticated ML techniques, in such conditions, addressing RQ1.

3.3. Crop model calibration scenarios

3.3.1. Intermediate crop model outputs: canopy cover and volumetric water content

Table 3 details performance indices computed between simulated and observed crop model outputs for each calibration scenario (Section 2.5.2.) for the five simulated plots taken together. Table A2 outlines the default values of all non-conservative parameters, which were either fixed using either secondary literature or observations, or calibrated under the different calibration scenarios (Section 2.5.2.). Fig. 6 illustrates the time series of the simulated and observed canopy cover (CC) (a-d) and volumetric water content (VWC) (e-h) for the same scenarios, as averages and standard error of the mean (SEM) for all the plots taken together. Tables 2 and A2, and Fig. 6 are used together to discuss the following results.

In the Default scenario, only crop growth parameters (i.e., CC_o, CC_x and CGC) were altered to improve the fit of the simulated CC curve with the observations. The SHPs which would impact the VWC curve fitting were chosen from (the average of) default values from literature (Table A2). CC was predicted very well as compared to the previous literature on AquaCrop wheat growth simulation, with r = 0.98 and RMSE = 5.71 % (Huang et al., 2022; Kale et al., 2018). These calibrated canopy growth parameters values were hence used in the $Raw_{Ref.}$, Raw_{LC} and Cal_{LC} scenarios, leading to nearly identical r and RMSE indicators. In all scenarios, the VWC simulations are compared against the observed soil moisture data from the reference sensor (Ref.).

For the VWC simulations, the Pearson r values did not show significant differences across the scenarios. However, when the SHPs were modified based on the raw Ref. and LC sensor values, the $Raw_{Ref.}$ (RMSE = 5.69 %) showed a significant improvement, and Raw_{LC} (RMSE = 8.58 %) showed a significant deterioration, both compared to the Default scenario (RMSE = 7.76 %). The worsening of the VWC simulation performance in the Raw_{LC} scenario in Fig. 6 (g), which overestimated the VWC_{obs} with a PBIAS = 46.19 %, resulted majorly from respectively higher VWC_{sim} values during the drying periods leading up to each of the irrigation days, when compared with the Default scenario in Fig. 6 (e), with a PBIAS = 27.74 %. This implies that the SHPs calibrated using raw LC soil moisture data were significantly poorer than those derived from the literature (Gupta et al., 2021; Rawls and Brakensiek, 1989), addressing research question (RQ) 2.

For the VWC simulations, the Pearson r values did not show

Table 3
Overall performance indices for major outputs of the AquaCrop-OS crop model - canopy cover (CC), volumetric water content (VWC) and biomass (BM), across the four crop model calibration scenarios (Default, Raw_{Ref}, Raw_{LC}, Cal_{LC}). The VWC simulations are compared against the observed soil moisture data from the reference sensor (Ref.). The effect of calibrating Soil Hydraulic Parameters (SHPs) on crop model performance of the crop model shows that calibrated LC sensor data ("Cal_{LC}" scenario) gives nearly as good results as raw Ref. data ("$Raw_{Ref.}$" scenario).

Crop model output	Performance Index	Default (Def)	$Raw_{Ref.}$ ($Raw_{Ref.}$)	Raw_{LC} (Raw_{LC})	$Calibrated_{LC}$ (Cal_{LC})
Scenario Descriptions		SHPs calibrated using default literature values	SHPs calibrated using raw VWC measurements from Ref. and LC sensor respectively	SHPs calibrated using 'best' LCcal model, 'PWLF-field'	
CC (%)	r (−)	0.98	0.98	0.98	0.98
	RMSE (%)	5.71	5.72	5.76	5.72
VWC (%)	r (−)	0.77	0.78	0.76	0.77
	RMSE (%)	7.76	5.70	9.55	5.90
	PBIAS (%)	27.74	3.99	46.19	5.22
BM (kg/m^2)	MAE (kg/m^2)	0.110	0.110	0.109	0.111
	RMSE (kg/m^2)	0.145	0.145	0.143	0.145
	NRMSE (%)	13.417	13.416	13.289	13.496
	PBIAS (%)	−0.63	−0.63	−0.57	−0.50

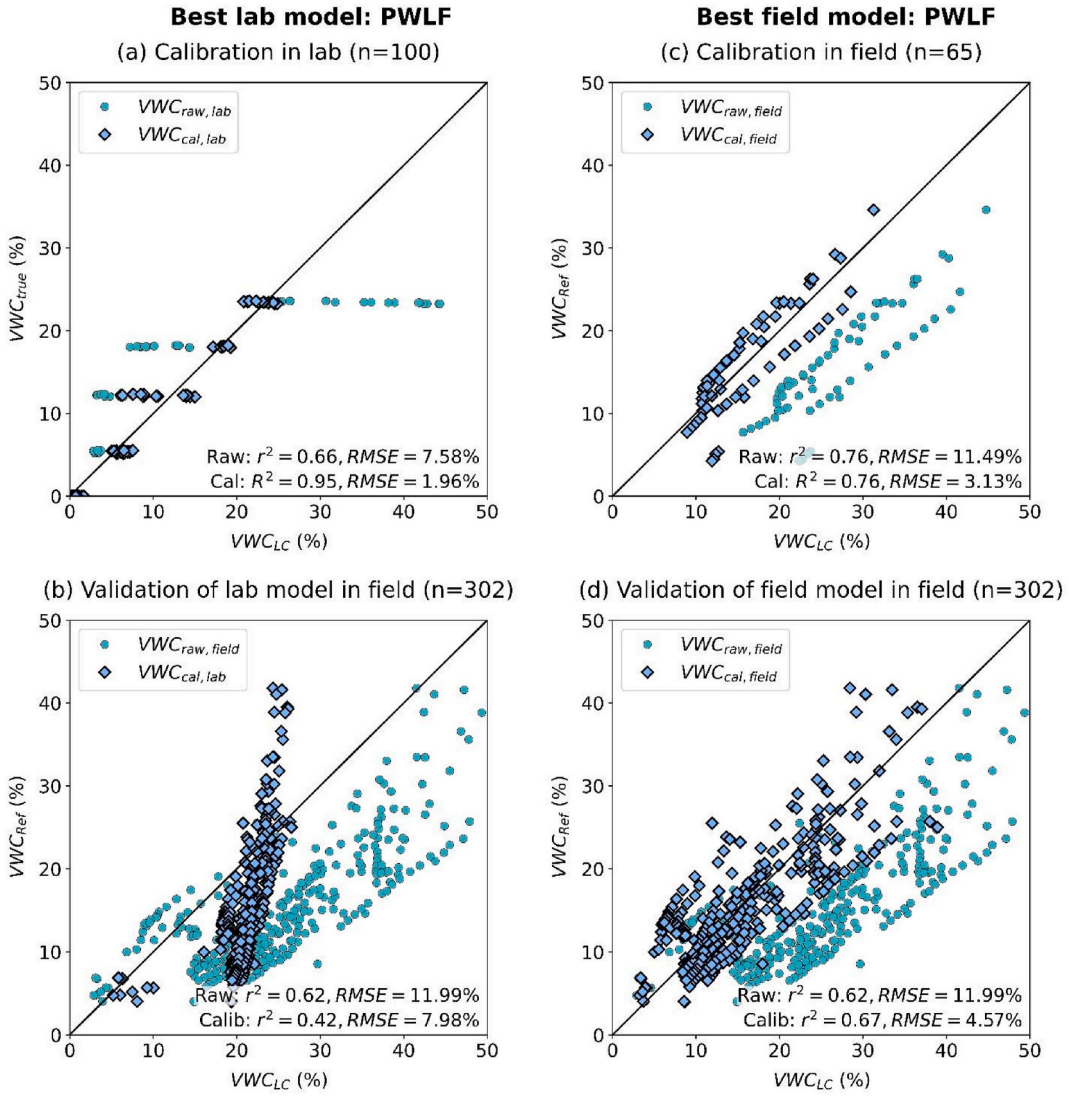

Fig. 5. Soil moisture calibration in the lab (top-left) and the field (top-right), validated on an independent, identical field dataset (bottom left and right respectively). r^2 represents the coefficient of determination. The best performing model in both cases is PWLF. The Y-axis represents VWC_{Ref} data, except 5(a) where it represents gravimetrically determined VWC.

significant differences across the scenarios. However, when the SHPs were modified based on the raw Ref. and LC sensor values, the $Raw_{Ref.}$ (RMSE = 5.69 %) showed a significant improvement, and Raw_{LC} (RMSE = 8.58 %) showed a significant deterioration, both compared to the Default scenario (RMSE = 7.76 %).

To answer research question (RQ) 2, the performances of the RAW_{LC}

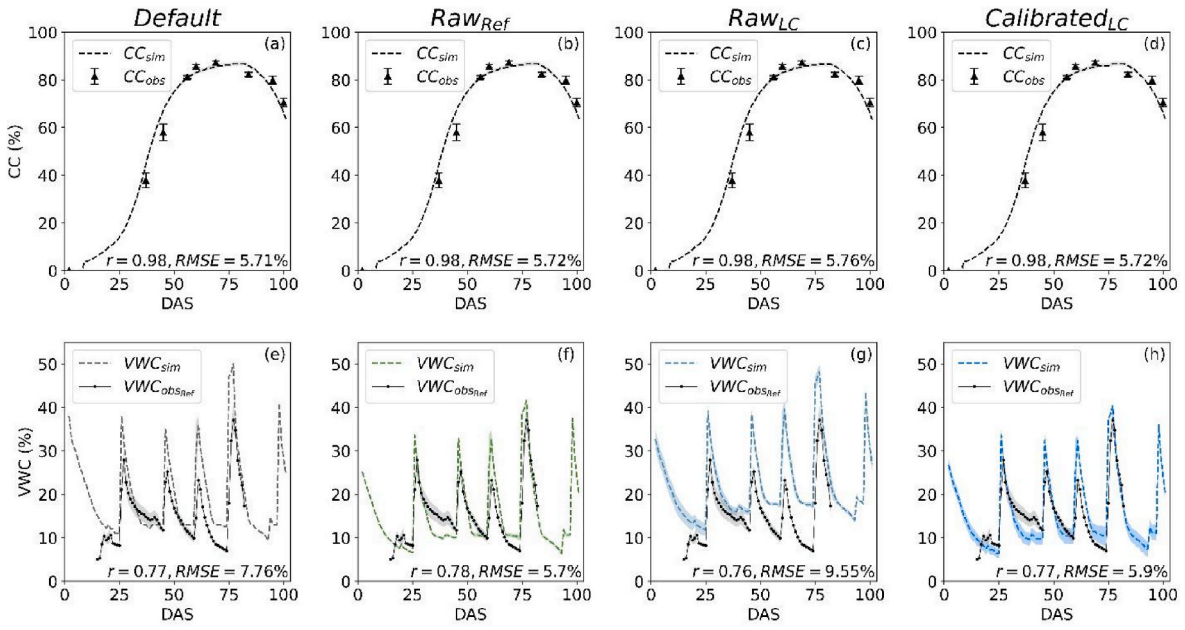

Fig. 6. Comparison of observed and AquaCrop-OS simulated ('sim') (a)–(d): Canopy cover (%) and (e)–(h): VWC (%) averaged over 5 field plots, for the 4 important AquaCrop-OS crop modeling scenarios (Default, $Raw_{Ref.}$, Raw_{LC} and Cal_{LC}). The VWC simulations are compared against the observed soil moisture data from the reference sensor (Ref.). Error bars and ranges in all figures represent the standard error of the respective mean values across the five validation plots. DAS implies days after sowing.

and Default scenarios were compared. The worsening of the VWC simulation performance in the Raw_{LC} scenario in Fig. 6 (g), which overestimated the VWC_{obs} with a PBIAS = 46.19 %, resulted majorly from respectively higher VWC_{sim} values during the drying periods leading up to each of the irrigation days, when compared with the Default scenario in Fig. 6 (e), with a PBIAS = 27.74 %. Hence, the SHPs calibrated using raw LC soil moisture data were significantly poorer than those derived from the literature (Gupta et al., 2021; Rawls and Brakensiek, 1989), implying the AquaCrop models calibrated with raw data from LC soil moisture performed relatively poorly compared to those calibrated using the literature, addressing RQ2.

The answer to RQ3 is that there was a significant difference between the VWC_{sim} performances of the SHPs calibrated using the Raw_{Ref} (RMSE = 5.7 %, PBIAS = 3.99 %) and Raw_{LC} (RMSE = 9.55 %, PBIAS = 46.19 %) data. This implies that the Raw_{Ref} derived SHPs led to significantly superior VWC_{sim} performances than that of the Raw_{LC} both in terms of accuracy and bias.

To answer RQ4, the SHPs were modified using the calibrated LC soil moisture data (in the Cal_{LC} scenario). The Cal_{LC} scenario (RMSE = 5.9 %, PBIAS = 5.22 %) in Fig. 6 (h) significantly outperformed the Raw_{LC} (RMSE = 9.55 %, PBIAS = 46.19 %), nearly attaining the performance of Raw_{Ref} (RMSE = 5.7 %, RMSE = 3.99 %). The SHPs calibrated in Cal_{LC} reduced the VWC_{sim} overestimation seen in the Raw_{LC} scenario. This was corroborated by the seasonal water balance computations, which showed that the SHPs interacted to reduce the final root zone soil water storage in the Cal_{LC} scenario to 19.85 % lower than the RAW_{LC} scenario (but still not as low as the $Raw_{Ref.}$ scenario). This implies that there were significant advantages of using LC soil moisture data calibrated using the Ref. sensor data, to calibrate the crop model SHPs, which addressed RQ4.

These results indicated the utility of calibrating LC sensors using Ref. sensors in terms of improving VWC simulation performance of AquaCrop, and that using raw LC sensor data would lead to poorer results than using the literature to calibrate SHPs.

However, the overall VWC simulation performance was not as good as previous studies which used either cumbersome gravimetric VWC measurements or costly neutron probes, both of which are highly accurate in VWC measurement (Huang et al., 2022; Kale et al., 2018; Zhang et al., 2013). Nevertheless, it was promising to note that significant improvements were obtained in the VWC simulations by calibrating the LC capacitance-based sensor even against an uncalibrated, reference (TDR-FDR) sensor.

3.3.2. *Final crop model outputs: crop model yield and water productivity*

Fig. 7 illustrates the simulated aboveground crop Water Productivity (WP_{ET}) for all the scenarios, as averages and standard error of the mean (SEM) for all the plots taken together. Table 3 also lists the MAE, RMSE and PBIAS for the respective biomass values. AquaCrop-OS overestimated the observed biomass by an average of 0.11 kg/m^2, and an average RMSE of 0.145 kg/m^2, and NRMSE of 13.41 %, which is considered to be a 'good' simulation performance (Jamieson et al., 1991; Raes et al., 2018). There were no significant differences in the simulated BM across the five scenarios irrespective of the strength of

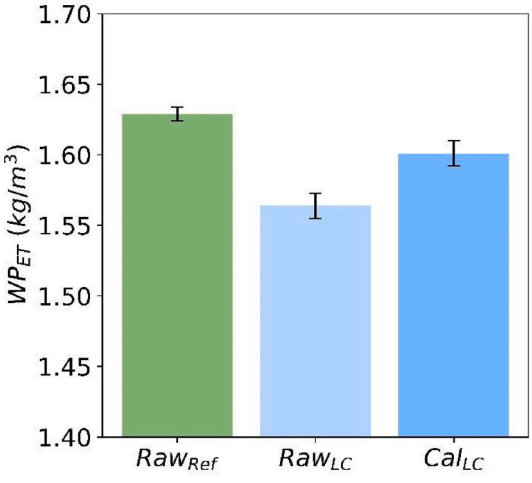

Fig. 7. Comparison of AquaCrop-OS simulated Water Productivity (WP_{ET}) averaged over 5 field plots, for the 3 different AquaCrop-OS crop modeling scenarios ($Raw_{Ref.}$, Raw_{LC} and $Calibrated_{LC}$). Error bars in all figures represent the SEM.

VWC simulation, indicating that the FAO AquaCrop simulation of biomass relied primarily on the canopy curve development.

However, the simulated water productivity (WP_{ET}) varied across the different scenarios. There was an improvement in WP_{ET} when the SHPs were calibrated using the raw Ref. sensor data as compared to the default literature (WP_{ET} for Def = 1.566 ± 0.009 kg/m^3, WP_{ET} for $Raw_{Ref.}$ = 1.629 ± 0.006 kg/m^3), due to significantly lower ET in $RAW_{Ref.}$ compared to Def. There was a non-significant change in WP_{ET} when raw LC sensor data were used to calibrate the SHPs: (WP_{ET} for Def = 1.566 ± 0.009 kg/m^3, WP_{ET} for Raw_{LC} = 1.564 ± 0.011 kg/m^3), which corresponded to the non-significant differences in simulated ET (and yield) between the two scenarios. This implies that incorporating the raw LC SM100 sensor data to calibrate the SHPs was not significantly more useful compared to using the literature values, and did not address any water saving objectives, further highlighting the ineffectiveness of using raw LC soil moisture data shown in Section 3.2.1.

However, when the LC sensors were calibrated using the Ref. sensors, there was a significant improvement in WP_{ET} (WP_{ET} for Raw_{LC} = 1.564 ± 0.011 kg/m^3, WP_{ET} for Cal_{LC} = 1.601 ± 0.010 kg/m^3), due to the significant differences between simulated ET between both the scenarios. This difference of 0.037 ± 0.021 kg/m^3 corresponds to water savings between 17,241 L and 62,500 L per kg of wheat yield produced.

3.4. Summary of results

This section summarizes the conclusions regarding each research question.

- The answer to RQ1 was that simpler models are more effective than sophisticated machine learning (ML) algorithms in calibrating LC soil moisture against data from a reference soil moisture sensor, since ML techniques can lead to model overfitting and overall less robust results. This was based on the discussion surrounding Table 2, and Figs. 4 and 5.
- The answer to RQ2 was that raw LC sensor data performed significantly poorly than literature values for parsimonious crop model calibration. There was no significant improvement in WP_{ET} compared to the default calibration case. These results were based on Fig. 6 in Section 3.3.1. However, sensor calibration may be difficult to implement due to the corresponding resource requirements (e.g., scientific, logistical, financial, human). If the LC sensors cannot be calibrated, using default values from the literature for SHP calibration can save sensor costs without compromising crop model performance.
- The answer to RQ3 was that the simulated AquaCrop VWC_{sim} output and WP_{ET} were poorer and significantly lower, respectively, when the raw LC soil moisture data was used to calibrate the SHPs (RMSE = 9.55 %, PBIAS = 46.19 %), compared to when the Reference soil moisture data were used (RMSE = 5.7 %, PBIAS = 3.99 %). These results were based on the discussion surrounding Fig. 6 in Section 3.3.1., and Section 3.3.2.
- The answer to RQ4 was that the effect of calibrating the raw LC soil moisture on the Reference sensor data significantly improved the VWC_{sim} output of the AquaCrop model compared to the respective output from the raw LC data, nearly attaining the performance based on the Reference sensor data. Additionally, this calibration led to significant improvements in simulated WP_{ET}. These results are based on the discussions in Section 3.3.1. (Figs. 6) and 3.3.2. (Fig. 7).

3.5. Further discussion

Capacitance based sensing is affected by environmental factors such as salinity and clay content, which can be measured and compensated for, to improve performance at lower frequencies of soil moisture sensing (Adla et al., 2020; Deng et al., 2020). The SM100 sensor has been shown to have low sensitivity to salinity by Adla et al. (2020), and their recommendations of soil specific calibrations have been followed in this study to account for soil textural variability. Similarly, the UGT manual (Umwelt-Geräte-Technik GmbH, 2017) mentions that the SMT100 sensor can operate well in clayey soils, and has low salinity sensitivity. Hence, this study implicitly accounts for these environmental factors.

The presence of a distinct wetting-front may also affect the calibration curves (Kargas et al., 2013; Young et al., 1997). While not capturing this phenomena is a limitation of the study, its impact on this particular application may have been relatively minor. The time for a wetting-front to pass the order of the sensor depth (~10 cm) (Chu et al., 2018) was much smaller than the time-resolution of the crop model (daily) for which soil moisture measurements were used. Also, the horizontal placement of the sensors was more likely to ensure uniform VWC within the sensing volume (Kargas et al., 2013).

Machine learning algorithms may have performed better by incorporating more data and variables (e.g., weather), but this was also not done to ensure comparability between the least squares and machine learning algorithms. The validations of the different soil moisture calibration models performed poorly than the recommended 3 % sensor accuracy for soil moisture sensor-based irrigation scheduling systems (Soulis et al., 2015). However, since this study was aimed at analyzing the effect of sensor calibration on relevant crop model outputs, this was not a valid concern within this context, and would need further investigation.

4. Conclusions

A set of best practices of calibrating capacitance based low-cost (LC) Spectrum SM100 soil moisture sensors was developed using two calibration approaches: calibrating against gravimetric water content in the using wet-up curve laboratory conditions, and calibrating against a TDR-FDR reference (Ref.) soil moisture sensor (UGT SMT100) in the field. Different least squares and machine learning approaches were used to calibrate the LC sensor. Field calibration was observed to be more robust than calibration in the laboratory. Calibration of the dry-down curve was found to be more accurate, and even calibrations using the dry-down curve in the laboratory were robust enough to perform similarly well as the superior field calibrations. Overfitting during calibration can lead to loss in robustness in the field, and hence, should be avoided. The best calibration model (LC_{cal}), considering the above issues, was the field calibrated piece-wise linear regression function (PWLF-field).

The soil hydraulic parameters (SHPs) of the FAO AquaCrop model were calibrated under different calibration schemes to understand the effect of using raw soil moisture sensor data, and the additional effect of calibrating the capacitance LC sensor, on model outputs. VWC estimation respectively improved and worsened on incorporating the Ref. and LC sensor data to calibrate SHPs, when compared with literature derived SHPs. However, LC data calibrated using the PWLF-field model resulted in nearly the same VWC simulation performance as the Ref. sensor. The water productivity (WP_{ET}) improved significantly when incorporating raw Ref. sensor data to calibrate SHPs. This was not seen when incorporating raw LC data, but on transforming the data using the best LC_{cal} model, a significant improvement was seen in the WP_{ET} compared to both the raw and default scenarios.

These experiments and modeling reveal that if calibration of LC sensors is not possible, it is preferable to use the literature to calibrate crop model SHPs without compromising overall model performance. However, calibrating the LC sensors using a higher-quality secondary standard sensor in the field may give rise to not only better VWC simulations by the crop model, but also significant improvements in water productivity and corresponding water savings at the same level of yield.

CRediT authorship contribution statement

Soham Adla: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software,

Validation, Visualization, Writing – original draft, Writing – review & editing. **Felix Bruckmaier:** Formal analysis, Investigation, Writing – review & editing, Writing – original draft, Methodology, Software, Data curation. **Leonardo F. Arias-Rodriguez:** Formal analysis, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Shivam Tripathi:** Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing, Funding acquisition. **Saket Pande:** Supervision, Writing – review & editing, Project administration. **Markus Disse:** Funding acquisition, Supervision, Writing – review & editing, Project administration, Resources.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Soham Adla reports financial support and travel were provided by Technical University of Munich Graduate School. Soham Adla reports administrative support was provided by Office of International Relations, Indian Institute of Technology Kanpur. Soham Adla reports financial support and equipment, drugs, or supplies were provided by Kritsnam Technologies Pvt. Ltd. Soham Adla reports equipment, drugs, or supplies was provided by Technical University of Munich Studienqualitaetskommission (SQK). Soham Adla reports equipment, drugs, or supplies was provided by Bayerische Landesantalt fuer Landwirtschaft (LfL). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The travel expenses to the Indian Institute of Technology (IIT) Kanpur were supported by the ‘ TUM-GS Internationalization Grant’ awarded by the Graduate School of the Technical University of Munich. Gratitude is expressed to the Office of International Relations, IIT Kanpur for its pleasant hospitality. The authors appreciate Kritsnam Technologies Pvt. Ltd. (in particular Mr. Neeraj Rai) and the staff of the Hydraulics Laboratory, Indian Institute of Technology Kanpur (specially Mr. Sunil Nishad) for their technical support throughout the measurement period. Also, the authors are grateful to Mr. Aditya Kumar, Mr. Pravesh Singh and Mr. Santosh for their overall field support for the experiment to be completed successfully. The Technische Universität München Studienqualitätskommission (SQK) 2017-18 funding was used to purchase the SMT100 soil moisture sensors deployed in the study. Soil testing for the study was conducted at the Bayerische Landesanstalt für Landwirtschaft (LfL), Germany.

Appendix

Details on the low-cost SM100 sensor

The capacitance-based Water Scout SM100 soil moisture sensor (Spectrum Technologies, Inc., Plainfield, IL, USA) operates with a pair of electrodes behaving as a capacitor and the soil surrounding the sensor behaving as the charge storing dielectric medium (Spectrum Technologies, 2014). An oscillator operating at 80 MHz drives the capacitor, and generates an output (voltage ratio) which is proportional to the dielectric permittivity of the soil-water system. The sensor output (voltage ratio/raw value) is then converted to a VWC value using the factory calibration equation - Equation (1) in the manuscript (Kieffer, 2018). The outputs of the SM100 sensor are "Raw value" (which is related to the voltage ratio of the output to the input voltage) and "VWC".

Details on the reference SMT100 sensor

The sensor head of the UGT SMT100 soil moisture sensor (manufactured by Umwelt-Geräte-Technik GmbH, Müncheberg, Germany) contains the sensor electronics which emit a steep pulse traveling along a closed transmission line buried in the soil. However, instead of directly measuring the time for the pulse to return (e.g., in TDT or TDR), the pulse is inverted and fed back into the line driver input, resulting in an "oscillation" frequency which is a function of the soil dielectric permittivity (Bogena et al., 2017). Also, unlike the FDR which is based on a capacitor, the SMT100 uses a ring oscillator to generate the pulse and transform the travel time to frequency (Bogena et al., 2017; Umwelt-Geräte-Technik GmbH, 2017). The oscillation frequency is around 340 MHz in air and 150 MHz in water (Bogena et al., 2017). This resultant frequency is high enough to not be influenced by the high clay content, electrical conductivity or imaginary dielectric permittivity of the soil (Blonquist et al., 2005; Umwelt-Geräte-Technik GmbH, 2017). The outputs of SMT100 are "Supply Voltage", "Count" (which can be linked to apparent dielectric permittivity through an empirical model, Bogena et al., 2017), "relative permittivity", "volumetric water content" (computed using the relative permittivity, Topp et al., 1980) and "soil temperature".

Table A1
Details of agricultural management during the experiment

Management Practice	Management Date	Management Date (DAS)	Management quantification (Irrigation depth (mm) or Quantity (kg/ha)
Fertilizer application	February 1, 2018	27 (at tillering)	347 kg/ha
Irrigation – 1	January 29, 2018	25	50
Irrigation – 2	February 18, 2018	45	43
Irrigation – 3	March 4, 2018	59	52.4
Irrigation – 4	March 19, 2018	74	37.3

DAS: Days after sowing.

Table A2
Non-conservative model parameters and their default values before calibrating the AquaCrop model. The target variable indicates which intermediate AquaCrop output is affected by a change in the corresponding parameter.

Parameter	Description	Target	Unit	Default value (or range)
SeedSize	Soil surface covered by an individual seedling at 90 % emergence	CC	cm^2/plant	1.5[(i)]
CCx	Maximum canopy cover	CC	$m^2\ m^{-2}$	0.80–0.99[(i)]
CGC	Canopy growth coefficient	CC	d^{-1} or °C-d^{-1}	0.0930–0.1235[(i)]
CDC	Canopy decline coefficient	CC	d^{-1} or °C-d^{-1}	0.0925[(i)]
Emergence	Time from sowing to emergence	CC	d	8
HIstart	Time from sowing to start of build-up of Harvest Index			70
Flowering	Length of the flowering stage	CC	d	14
Senescence	Time from sowing to start of senescence	CC	d	81
Maturity	Time from sowing to maturity, i.e., length of crop cycle	CC	d	100
Zmin	Minimum effective rooting depth	VWC	m	0.2–0.3[(i)]
Zmax	Maximum effective rooting depth	VWC	m	1.5
AppEff	Irrigation Application Efficiency	VWC	%	60[(ii, iii)]
REW	Readily Evaporable Water	VWC	mm	9–12[(iv)]
K_sat	Saturated hydraulic conductivity	VWC	mm-d^{-1}	96-446[(v, vi)]
th_wp	VWC at Permanent wilting point	VWC	$m^3\ m^{-3}$	0.133 (0.078–0.188)[(v)]
th_fc	VWC at Field Capacity	VWC	$m^3\ m^{-3}$	0.330 (0.258–0.402)[(v)]
th_sat	VWC at Saturation	VWC	$m^3\ m^{-3}$	0.501 (0.42–0.582)[(v)]
CN	Curve Number for antecedent moisture class II	VWC	–	69–75[(vii)]
HI	Reference Harvest Index	Yield	%	0.48 (0.45–0.50)[(i)]

(i) Raes et al. (2018), (ii) Brouwer et al. (1989), (iii) Taghavaeian (2017), (iv) Allen et al. (2005), (v) Rawls and Brakensiek (1989), (vi) Gupta et al. (2021), (vii) USDA-NRCS (2017).

Table A3
Fitted hyperparameter sets used in each algorithm for calibrating soil moisture data.

Model	Hyperparameters
Linear Regression (LR)	fit_intercept = True, copy_X = True, n_jobs = None, positive = False
Support Vector Regression (SVR)	kernel = 'rbf', degree = 3, gamma = 'scale', coef0 = 0.0, tol = 0.001, C = 1.0, epsilon = 0.1, shrinking = True, cache_size = 200, verbose = False, max_iter = −1
Random Forest Regressor (RFR)	n_estimators = 100, *, criterion = 'squared_error', max_depth = None, min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, max_features = 1.0, max_leaf_nodes = None, min_impurity_decrease = 0.0, bootstrap = True, oob_score = False, n_jobs = None, random_state = None, verbose = 0, warm_start = False, ccp_alpha = 0.0, max_samples = None
Multi Layer Perception (MLP)	hidden_layer_sizes=(100), activation = 'relu', *, solver = 'adam', alpha = 0.0001, batch_size = 'auto', learning_rate = 'constant', learning_rate_init = 0.001, power_t = 0.5, max_iter = 200, shuffle = True, random_state = None, tol = 0.0001, verbose = False, warm_start = False, momentum = 0.9, nesterovs_momentum = True, early_stopping = False, validation_fraction = 0.1, beta_1 = 0.9, beta_2 = 0.999, epsilon = 1e-08, n_iter_no_change = 10, max_fun = 15,000

References

Adla, S., Rai, N.K., Karumanchi, S.H., Tripathi, S., Disse, M., Pande, S., 2020. Laboratory calibration and performance evaluation of low-cost capacitive and very low-cost resistive soil moisture sensors. Sensors 20, 363.

Ahmed, M., Akram, M., Asim, M., Aslam, M., Hassan, F., Higgins, S., Stöckle, C., Hoogenboom, G., 2016. Calibration and validation of APSIM-Wheat and CERES-wheat for spring wheat under rainfed conditions: models evaluation and application. Comput. Electron. Agric. 123, 384–401. https://doi.org/10.1016/j.compag.2016.03.015.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56. Food and Agricultural Organization of the United Nations, Rome.

Allen, R.G., Pereira, L.S., Smith, M., Raes, D., Wright, J.L., 2005. FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions. J. Irrigat. Drain. Eng. 131, 2–13. /(ASCE)0733-9437(2005)131:1(2).

Andarzian, B., Bannayan, M., Steduto, P., Mazraeh, H., Barati, M.E., Barati, M.A., Rahnama, A., 2011. Validation and testing of the AquaCrop model under full and deficit irrigated wheat production in Iran. Agric. Water Manag. 100, 1–8.

AquaCrop, 2016. Tutorials. Italy, Rome.

Arias-Rodriguez, L.F., Duan, Z., Díaz-Torres, J.d.J., Basilio Hazas, M., Huang, J., Kumar, B.U., Tuo, Y., Disse, M., 2021. Integration of remote sensing and mexican water quality monitoring system using an extreme learning machine. Sensors [Online] 21 (12), 4118.doi. https://doi.org/10.3390/s21124118.

Asseng, S., Zhu, Y., Basso, B., Wilson, T., Cammarano, D., 2014. Simulation modeling: applications in cropping systems. In: Van Alfen, N. (Ed.), Encyclopedia of Agriculture and Food Systems. Academic Press, Oxford, pp. 102–112.

Attia, A., Rajan, N., Xue, Q., Nair, S., Ibrahim, A., Hays, D., 2016. Application of DSSAT-CERES-Wheat model to simulate winter wheat response to irrigation management in the Texas High Plains. Agric. Water Manag. 165, 50–60. https://doi.org/10.1016/j.agwat.2015.11.002.

Babel, M., Deb, P., Soni, P., 2019. Performance evaluation of AquaCrop and DSSAT-CERES for maize under different irrigation and manure application rates in the himalayan region of India. Agric. Res. 8, 207–217. https://doi.org/10.1007/s40003-018-0366-y.

Bastiaanssen, W.G., Molden, D.J., Makin, I.W., 2000. Remote sensing for irrigated agriculture: examples from research and possible applications. Agric. Water Manag. 46, 137–155.

Bello, Z.A., Tfwala, C.M., van Rensburg, L.D., 2019. Evaluation of newly developed capacitance probes for continuous soil water measurement. Geoderma 345, 104–113. https://doi.org/10.1016/j.geoderma.2019.03.030.

Bjorneberg, D., 2013. Irrigation Methods, Reference Module in Earth Systems and Environmental Sciences. Elsevier Inc.

Blonquist, J.M., Jones, S.B., Robinson, D.A., 2005. Standardizing characterization of electromagnetic water content sensors: Part 2. Evaluation of seven sensing systems. Vadose Zone J. 4, 1059–1069. https://doi.org/10.2136/vzj2004.0141.

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

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brouwer, C., Prins, K., Heibloem, M., 1989. Annex I: irrigation efficiencies. In: Irrigation Water Management: Irrigation Scheduling. Food and Agricultural Organization of the United Nations, Rome, Italy.

Burns, T.T., Adams, J.R., Berg, A.A., 2014. Laboratory calibration procedures of the Hydra probe soil moisture sensor:infiltration wet-up vs. Dry-down. Vadose Zone J. 13 https://doi.org/10.2136/vzj2014.07.0081 vzj2014.07.0081.

Carranza, C., Nolet, C., Pezij, M., van der Ploeg, M., 2021. Root zone soil moisture estimation with Random Forest. J. Hydrol. 593, 125840.

Chai, S.-S., Walker, J.P., Makarynskyy, O., Kuhn, M., Veenendaal, B., West, G., 2009. Use of soil moisture variability in artificial neural network retrieval of soil moisture. Rem. Sens. 2, 166–190.

Chanasyk, D., Naeth, M., 1996. Field measurement of soil moisture using neutron probes. Can. J. Soil Sci. 76, 317–323. https://doi.org/10.4141/cjss96-038.

Chandler, D.G., Seyfried, M., Murdock, M., McNamara, J.P., 2004. Field calibration of water content reflectometers. Soil Sci. Soc. Am. J. 68, 1501–1507. https://doi.org/10.2136/sssaj2004.1501.

Chen, C., Wang, E., Yu, Q., 2010. Modelling the effects of climate variability and water management on crop water productivity and water balance in the North China Plain. Agric. Water Manag. 97, 1175–1184. https://doi.org/10.1016/j.agwat.2008.11.012.
Chenu, K., Chapman, S., Tardieu, F., McLean, G., Welcker, C., 2009. Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: a "gene-to-phenotype" modeling approach. Genetics 183, 1507–1523. https://doi.org/10.1534/genetics.109.105429.
Chiara, C., Marco, M., 2022. Irrigation efficiency optimization at multiple stakeholders' levels based on remote sensing data and energy water balance modelling. Irrigat. Sci. https://doi.org/10.1007/s00271-022-00780-4.
Chu, X., Jia, X., Liu, Y., 2018. Quantification of wetting front movement under the influence of surface topography. Soil Res. 56, 382–395. https://doi.org/10.1071/SR17071.
Cosh, M., Jackson, T., Bindlish, R., Famiglietti, J., Ryu, D., 2005. Calibration of an impedance probe for estimation of surface soil water content over large regions. J. Hydrol. 311, 49–58. https://doi.org/10.1016/j.jhydrol.2005.01.003.
Cummings, R., Chandler Jr., R., 1941. A field comparison of the electrothermal and gypsum block electrical resistance methods with the tensiometer method for estimating soil moisture in situ. Soil Sci. Soc. Am. J. 5, 80–85. https://doi.org/10.2136/sssaj1941.036159950005000C0015x.
Deng, X., Gu, H., Yang, L., Lyu, H., Cheng, Y., Pan, L., Fu, Z., Cui, L., Zhang, L., 2020. A method of electrical conductivity compensation in a low-cost soil moisture sensing measurement based on capacitance. Measurement 150, 107052. https://doi.org/10.1016/j.measurement.2019.107052.
Doorenbos, J., Kassam, A.H., 1979. Yield response to water. (FAO Irrigation and Drainage Paper No. Food and Agricultural Organization of the United Nations.
Dwivedi, S.K., Kumar, S., Mishra, J.S., Haris, A.A., Singh, S.K., Srivastava, A.K., Kumar, A., Kumar, V., Singh, S., Bhatt, B.P., 2019. Effect of moisture regimes and sowing dates on wheat physiological process and yield attributes under rain-fed ecosystem in Eastern Indo Gangetic Plain. Plant Physiol. Rep. 24, 46–53. https://doi.org/10.1007/s40502-018-0406-4.
Eitzinger, J., Trnka, M., Hösch, J., Žalud, Z., Dubrovský, M., 2004. Comparison of CERES, WOFOST and SWAP models in simulating soil water content during growing season under different soil conditions. Ecol. Model. 171, 223–246.
Elshorbagy, A., Corzo, G., Srinivasulu, S., Solomatine, D., 2010. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology-Part 1: concepts and methodology. Hydrol. Earth Syst. Sci. 14, 1931–1941.
FAO, 2016. Retrieved from: https://www.fao.org/aquacrop/resources/tutorials/en/.
Fares, A., Alva, A., 2000. Evaluation of capacitance probes for optimal irrigation of citrus through soil moisture monitoring in an entisol profile. Irrigat. Sci. 19, 57–64. https://doi.org/10.1007/s002710050001.
Feng, G., Sui, R., 2020. Evaluation and calibration of soil moisture sensors in undisturbed soils. Transactions of the ASABE 63, 265–274. https://doi.org/10.13031/trans.13428.
Foster, T., Brozović, N., Butler, A., Neale, C., Raes, D., Steduto, P., Fereres, E., Hsiao, T.C., 2017. AquaCrop-OS: an open source version of FAO's crop water productivity model. Agric. Water Manag. 181, 18–22.
Gallardo, M., Elia, A., Thompson, R.B., 2020. Decision support systems and models for aiding irrigation and nutrient management of vegetable crops. Agric. Water Manag. 240 https://doi.org/10.1016/j.agwat.2020.106209.
García, G., Brogioni, M., Venturini, V., Rodriguez, L., Fontanelli, G., Walker, E., Graciani, S., Macelloni, G., 2016. Soil moisture estimation using multi linear regression with terraSAR-X data. Revista de Teledetección 73–81.
Gaskin, G., Miller, J., 1996. Measurement of soil water content using a simplified impedance measuring technique. J. Agric. Eng. Res. 63, 153–159. https://doi.org/10.1006/jaer.1996.0017.
Gedilu, E., 2020. Soil Moisture Dynamics in Agricultural Fields with Heterogeneous Unsaturated Soils (PhD). Indian Institute of Technology Kanpur, Kanpur.
Gill, M.K., Asefa, T., Kemblowski, M.W., McKee, M., 2006. Soil moisture prediction using support vector machines 1. JAWRA Journal of the American Water Resources Association 42, 1033–1046.
Graves, A., Hess, T., Matthews, R., 2002. Using models as tools in education and training. Crop-soil simulation models: applications in developing countries 151–182.
Gu, Z., Zhu, T., Jiao, X., Xu, J., Qi, Z., 2021. Neural network soil moisture model for irrigation scheduling. Comput. Electron. Agric. 180, 105801.
Gupta, S., Sri Harsha, K., Dash, S.K., Adla, S., Tripathi, S., Sinha, R., Paul, D., Sen, I.S., 2019. Monitoring ecosystem health in India's food basket. Eos 100. https://doi.org/10.1029/2019EO117683.
Gupta, S., Hengl, T., Lehmann, P., Bonetti, S., Or, D., 2021. SoilKsatDB: global database of soil saturated hydraulic conductivity measurements for geoscience applications. Earth Syst. Sci. Data 13, 1593–1612.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, second ed. Springer Series in Statistics.
Huang, M., Wang, C., Qi, W., Zhang, Z., Xu, H., 2022. Modelling the integrated strategies of deficit irrigation, nitrogen fertilization, and biochar addition for winter wheat by AquaCrop based on a two-year field study. Field Crops Res. 282, 108510.
Iqbal, M.A., Shen, Y., Stricevic, R., Pei, H., Sun, H., Amiri, E., Penas, A., del Rio, S., 2014. Evaluation of the FAO AquaCrop model for winter wheat on the North China Plain under deficit irrigation from field experiment to regional yield simulation. Agric. Water Manag. 135, 61–72.
Jamieson, P., Porter, J., Wilson, D., 1991. A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. Field Crops Res. 27, 337–350.
Jekel, C., Venter, G., 2019. pwlf: A Python Library for Fitting 1D Continuous Piecewise Linear Functions. Available at: https://github.com/cjekel/piecewise_linear_fit_py.
Jones, S., Blonquist Jr., J., Robinson, D., Rasmussen Or, D., 2005. Standardizing characterization of electromagnetic water content sensors: Part 1. Methodology. Vadose Zone J. 4, 1048–1058. https://doi.org/10.2136/vzj2004.0140.
Jones, J.W., Naab, J., Fatondji, D., Dzotsi, K., Adiku, S., He, J., 2012. Uncertainties in simulating crop performance in degraded soils and low input production systems. In: Improving Soil Fertility Recommendations in Africa Using the Decision Support System for Agrotechnology Transfer (DSSAT). Springer, pp. 43–59.
Kale, C., Madenoglu, S., Sonmez, B., 2018. Evaluating AquaCrop model for winter wheat under various irrigation conditions in Turkey. Tarim Bilimleri Dergisi 24, 205–217.
Kargas, G., Soulis, K.X., 2012. Performance analysis and calibration of a new low-cost capacitance soil moisture sensor. J. Irrigat. Drain. Eng. 138, 632–641.
Kargas, G., Soulis, K.X., 2019. Performance evaluation of a recently developed soil water content, dielectric permittivity, and bulk electrical conductivity electromagnetic sensor. Agric. Water Manag. 213, 568–579.
Kargas, G., Ntoulas, N., Nektarios, P.A., 2013. Soil texture and salinity effects on calibration of TDR300 dielectric moisture sensor. Soil Res. 51, 330–340. https://doi.org/10.1071/SR13009.
Keating, B., Carberry, P., Hammer, G., Probert, M., Robertson, M., Holzworth, D., Huth, N., Hargreaves, J., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J., Silburn, M., Wang, E., Brown, S., Bristow, K., Asseng, S., Chapman, S., McCown, R., Freebairn, D., Smith, C., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18, 267–288. https://doi.org/10.1016/S1161-0301(02)00108-9.
Kieffer, D., 2018. Personal Communication.
Kisekka, I., Peddinti, S.R., Kustas, W.P., McElrone, A.J., Bambach-Ortiz, N., McKee, L., Bastiaanssen, W., 2022. Spatial–temporal modeling of root zone soil moisture dynamics in a vineyard using machine learning and remote sensing. Irrigat. Sci. https://doi.org/10.1007/s00271-022-00775-1.
König, N., Blum, U., Symossek, F., 2005. Handbuch forstliche analytik (HFA. In:). Eine Loseblatt-Sammlung der Analysemethoden im Forstbereich (5. Ergänzung 2014).
Kumar, A., 2019. Partitioning of Evapotranspiration Using Hydrometric Methods (Master's Thesis). Indian Institute of Technology Kanpur, Kanpur.
Kumar, S., Sharma, V., Chaudhary, S., Tyagi, A., Mishra, P., Priyadarshini, A., Singh, A., 2012. Genetics of flowering time in bread wheat Triticum aestivum: complementary interaction between vernalization-insensitive and photoperiod-insensitive mutations imparts very early flowering habit to spring wheat. J. Genet. 91, 33–47. https://doi.org/10.1007/s12041-012-0149-3.
Kumar, S., Dwivedi, S.K., Kumar, R., Mishra, J.S., Singh, S.K., Prakash, V., Rao, K.K., Bhatt, B.P., 2017. Productivity and energy use efficiency of wheat (Triticum aestivum) genotypes under different tillage options in rainfed ecosystem of middle Indo-Gangetic Plains. Indian J. Agron. 62, 31–38.
Landau, S., Mitchell, R., Barnett, V., Colls, J., Craigon, J., Payne, R., 2000. A parsimonious, multiple-regression model of wheat yield response to environment. Agric. For. Meteorol. 101, 151–166. https://doi.org/10.1016/S0168-1923(99)00166-5.
Lee, Y., Jung, C., Kim, S., 2019. Spatial distribution of soil moisture estimates using a multiple linear regression model and Korean geostationary satellite (COMS) data. Agric. Water Manag. 213, 580–593.
Liang, H., Qi, Z., DeJonge, K.C., Hu, K., Li, B., 2017. Global sensitivity and uncertainty analysis of nitrate leaching and crop yield simulation under different water and nitrogen management practices. Comput. Electr. Agri. 142, 201–210. https://doi.org/10.1016/j.compag.2017.09.010.
Liu, D., Yu, Z., Hai-shen, L., 2010. Data assimilation using support vector machines and ensemble Kalman filter for multi-layer soil moisture prediction. Water Sci. Eng. 3, 361–377.
Lu, Y., Chibarabada, T.P., Ziliani, M.G., Onema, J.-M.K., McCabe, M.F., Sheffield, J., 2021. Assimilation of soil moisture and canopy cover data improves maize simulation using an under-calibrated crop model. Agric. Water Manag. 252, 106884.
Ma, H., Malone, R.W., Jiang, T., Yao, N., Chen, S., Song, L., Feng, H., Yu, Q., He, J., 2020. Estimating crop genetic parameters for DSSAT with modified PEST software. Euro. J. Agro. 115, 126017. https://doi.org/10.1016/j.eja.2020.126017.
Matula, S., Bát'ková, K., Legese, W.L., 2016. Laboratory performance of five selected soil moisture sensors applying factory and own calibration equations for two soil media of different bulk density and salinity levels. Sensors 16, 1912.
Muñoz-Carpena, R., Yuncong, C., Klassen, W., Dukes, M., 2005. Field comparison of tensiometer and granular matrix sensor automatic drip irrigation on tomato. HortTechnology 15. https://doi.org/10.21273/HORTTECH.15.3.0584.
Nagahage, E.A.A.D., Nagahage, I.S.P., Fujino, T., 2019. Calibration and validation of a low-cost capacitive moisture sensor to integrate the automated soil moisture monitoring system. Agriculture 9, 141. https://doi.org/10.3390/agriculture9070141.
Nakra, B.C., Chaudhry, K.K., 2006. Introduction to Instruments and Their Representation. In: Instrumentation, Measurement and. Analysis.
Navarro-Hellín, H., Torres-Sánchez, R., Soto-Valles, F., Albaladejo-Pérez, C., López-Riquelme, J.A., Domingo-Miguel, R., 2015. A wireless sensors architecture for efficient irrigation water management. Agric. Water Manag. 151, 64–74.
Nielsen, D.C., Miceli-Garcia, J.J., Lyon, D.J., 2012. Canopy cover and leaf area index relationships for wheat, triticale, and corn. Agron. J. 104, 1569–1573.
Ojo, E.R., Bullock, P.R., L'Heureux, J., Powers, J., McNairn, H., Pacheco, A., 2015. Calibration and evaluation of a frequency domain reflectometry sensor for real-time soil moisture monitoring. Vadose Zone J. 14, 1–12. https://doi.org/10.2136/vzj2014.08.0114.
Peddinti, S.R., Hopmans, J.W., Najm, M.A., Kisekka, I., 2020. Assessing effects of salinity on the performance of a low-cost wireless soil water sensor. Sensors 20. https://doi.org/10.3390/s20247041.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Placidi, P., Gasperini, L., Grassi, A., Cecconi, M., Scorzoni, A., 2020. Characterization of low-cost capacitive soil moisture sensors for IoT networks. Sensors 20, 3585. https://doi.org/10.3390/s20123585.
Placidi, P., Morbidelli, R., Fortunati, D., Papini, N., Gobbi, F., Scorzoni, A., 2021. Monitoring soil and ambient parameters in the IoT precision agriculture scenario: an original modeling approach dedicated to low-cost soil water content sensors. Sensors 21, 5110. https://doi.org/10.3390/s21155110.
Pramanik, M., Khanna, M., Singh, M., Singh, D.K., Sudhishri, S., Bhatia, A., Ranjan, R., 2022. Automation of soil moisture sensor-based basin irrigation system. Smart Agricultural Technology 2, 100032. https://doi.org/10.1016/j.atech.2021.100032.
Qingling, B., Jianli, D., Jingzhe, W., 2019. Prediction of soil moisture content by selecting spectral characteristics using random forest method. Laser & Optoelectronics Progress 55, 113002.
Qiu, Y., Fu, B., Wang, J., Chen, L., 2003. Spatiotemporal prediction of soil moisture content using multiple-linear regression in a small catchment of the Loess Plateau, China. Catena 54, 173–195.
Quintero, D., Díaz, E., 2020. A comparison of two open-source crop simulation models for a potato crop. Agron. Colomb. 38 https://doi.org/10.15446/agron.colomb.v38n3.82525.
Raes, D., 2017. AquaCrop Training Handbook I. Understanding AquaCrop, FAO. (Training Handbook). Food and Agricultural Organization of the United Nations, Rome, Italy.
Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2018. AquaCrop Version 6.0 – 6.1, Reference Manual (Reference Manual). Food and Agricultural Organization of the United Nations, Rome, Italy.
Rai, P., 2012. Calibration of Soil Moisture Capacitance Probe (Master). Indian Institute of Technology Kanpur, Kanpur.
Raine, S.R., 2006. The Future of Irrigation Technologies and Practices Is Here Now. Keynote Address.
Rawls, W.J., Brakensiek, D.L., 1989. Estimation of soil water retention and hydraulic properties. In: Unsaturated Flow in Hydrologic Modeling. Springer, Dordrecht, pp. 275–300. https://doi.org/10.1007/978-94-009-2352-2_10.
Robinson, D., Jones, S.B., Wraith, J., Or, D., Friedman, S., 2003. A review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry. Vadose Zone J. 2, 444–475. https://doi.org/10.2113/2.4.444.
Rodríguez-Robles, J., Martin, Á., Martin, S., Ruipérez-Valiente, J.A., Castro, M., 2020. Autonomous sensor network for rural agriculture environments, low cost, and energy self-charge. Sustainability 12, 5913.
Rosenbaum, U., Huisman, J.A., Weuthen, A., Vereecken, H., Bogena, H.R., 2010. Sensor-to-Sensor variability of the ECH2O EC-5, TE, and 5TE sensors in dielectric liquids. Vadose Zone J. 9, 181–186. https://doi.org/10.2136/vzj2009.0036.
Rosenbaum, U., Huisman, J.A., Vrba, J., Vereecken, H., Bogena, H.R., 2011. Correction of temperature and electrical conductivity effects on dielectric permittivity measurements with ECH2O sensors. Vadose Zone J. 10, 582–593. https://doi.org/10.2136/vzj2010.0083.
Rudnick, D.R., Djaman, K., Irmak, 2015. Performance analysis of capacitance and electrical resistance-type soil moisture sensors in a silt loam soil. Trans. ASABE Am. Soc. Agric. Biol. Eng. 649–665. https://doi.org/10.13031/trans.58.10761.
Sankararamakrishnan, N., Sharma, A.K., Iyengar, L., 2008. Contamination of nitrate and fluoride in ground water along the Ganges Alluvial Plain of Kanpur district, Uttar Pradesh, India. Environ. Monit. Assess. 146, 375–382. https://doi.org/10.1007/s10661-007-0085-5.
Saseendran, S., Trout, T., Ahuja, L., Ma, L., McMaster, G., Nielsen, D., Andales, A., Chávez, J., Ham, J., 2015. Quantifying crop water stress factors from soil water measurements in a limited irrigation experiment. Agric. Syst. 137, 191–205.
Silva, J.V., Giller, K.E., 2020. Grand challenges for the 21st century: what crop models can and can't (yet) do. J. Agric. Sci. 158, 794–805. https://doi.org/10.1017/S0021859621000150.
Silva, J.V., Reidsma, P., Laborte, A., van Ittersum, M., 2017. Explaining rice yields and yield gaps in Central Luzon, Philippines: an application of stochastic frontier analysis and crop modelling. Eur. J. Agron. 82, 223–241.
Singh, M., 2014. Techniques of crop cutting experiment. In: Recent Advances in Survey Design and Analysis of Survey Data Using Statistical Software, Reference Manual. ICAR Indian Agricultural Statistics Research Institute, New Delhi.
Singh, A.K., Rahman, A., Sharma, S.P., Upadhyaya, A., Sikka, A.K., 2009. Small holders' irrigation—problems and options. Water Resour. Manag. 23, 289–302. https://doi.org/10.1007/s11269-008-9275-3.
Singh, A., Saha, S., Mondal, S., 2013. Modelling irrigated wheat production using the FAO AquaCrop model in West Bengal, India, for sustainable agriculture. Irrigat. Drain. 62, 50–56.
Singh, J., Lo, T., Rudnick, D.R., Dorr, T.J., Burr, C.A., Werle, R., Shaver, T.M., Muñoz-Arriola, F., 2018. Performance assessment of factory and field calibrations for electromagnetic sensors in a loam soil. Agric. Water Manag. 196, 87–98. https://doi.org/10.1016/j.agwat.2017.10.020.
Soltani, A., Sinclair, T., 2015. A comparison of four wheat models with respect to robustness and transparency: simulation in a temperate, sub-humid environment. Field Crops Res. 175, 37–46.
Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of rainfall-runoff models: application of global optimization to the sacramento soil moisture accounting model. Water Resour. Res. 29, 1185–1194.
Soulis, K.X., Elmaloglou, S., Dercas, N., 2015. Investigating the effects of soil moisture sensors positioning and accuracy on soil moisture based drip irrigation scheduling systems. Agric. Water Manag. 148, 258–268.
Spectrum Technologies, 2014. WaterScout SM100 Soil Moisture Sensor Product Manual. Spectrum Technologies, Inc., 12350 S. Industrial Dr. East Plainfield, , IL (USA).
Spelman, D., Kinzli, K.-D., Kunberger, T., 2013. Calibration of the 10HS soil moisture sensor for southwest Florida agricultural soils. J. Irrigat. Drain. Eng. 139 https://doi.org/10.1061/(ASCE)IR.1943-4774.0000647.
Srbinovska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., Borozan, V., 2015. Environmental parameters monitoring in precision agriculture using wireless sensor networks. J. Clean. Prod. 88, 297–307. https://doi.org/10.1016/j.jclepro.2014.04.036.
Srivastava, P.K., Petropoulos, G.P., Prasad, R., Triantakonstantis, D., 2021. Random forests with bagging and genetic algorithms coupled with least trimmed squares regression for soil moisture deficit using SMOS satellite soil moisture. ISPRS Int. J. Geo-Inf. 10, 507.
Steduto, P., Hsiao, T.C., Raes, D., Fereres, E., 2009. AquaCrop—the FAO crop model to simulate yield response to water: I. Concepts and underlying principles. Agron. J. 101, 426–437. https://doi.org/10.2134/agronj2008.0139s.
Steduto, P., Hsiao, T.C., Fereres, E., Raes, D., 2012. Crop Yield Response to Water (No. 66). FAO Irrigation and Drainage Paper. Food and Agricultural Organization of the United Nations, Rome, Italy.
Taghavaeian, S., 2017. Surface Irrigation Systems (Fact Sheet No. BAE-1527), Oklahoma Cooperative Extension Fact Sheets. Oklahoma State University, Oklahoma, USA.
Teng, W., Wang, J., Doraiswamy, P., 1993. Relationship between satellite microwave radiometric data, antecedent precipitation index, and regional soil moisture. Int. J. Rem. Sens. 14, 2483–2500.
Thompson, R.B., Gallardo, M., Valdez, L.C., Fernández, M.D., 2007a. Using plant water status to define threshold values for irrigation management of vegetable crops using soil moisture sensors. Agric. Water Manag. 88, 147–158. https://doi.org/10.1016/j.agwat.2006.10.007.
Thompson, R.B., Gallardo, M., Valdez, L.C., Fernández, M.D., 2007b. Determination of lower limits for irrigation management using in situ assessments of apparent crop water uptake made with volumetric soil water content sensors. Agric. Water Manag. 92, 13–28. https://doi.org/10.1016/j.agwat.2007.04.009.
Todorovic, M., Albrizio, R., Zivotic, L., Saab, M., Stöckle, C., Steduto, P., 2009. Assessment of AquaCrop, CropSyst, and WOFOST models in the simulation of sunflower growth under different water regimes. Agron. J. 101, 509–521. https://doi.org/10.2134/agronj2008.0166s.
Topp, G.C., Davis, J.L., Annan, A.P., 1980. Electromagnetic determination of soil water content: measurements in coaxial transmission lines. Water Resour. Res. 16, 574–582. https://doi.org/10.1029/WR016i003p00574.
Toumi, J., Er-Raki, S., Ezzahar, J., Khabba, S., Jarlan, L., Chehbouni, A., 2016. Performance assessment of AquaCrop model for estimating evapotranspiration, soil water content and grain yield of winter wheat in Tensift Al Haouz (Morocco): application to irrigation management. Agric. Water Manag. 163, 219–235.
Umwelt-Geräte-Technik GmbH, 2017. 01/2017/EN Catalog. Umwelt-Geräte-Technik GmbH.
USDA-NRCS, 2017. Hydrologic soil-cover complexes. In: National Engineering Handbook. United States Department of Agriculture.
Van Halsema, G.E., Vincent, L., 2012. Efficiency and productivity terms for water management: a matter of contextual relativism versus general absolutism. Agric. Water Manag. 108, 9–15. https://doi.org/10.1016/j.agwat.2011.05.016.
Vanuytrecht, E., Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., Heng, L.K., Vila, M.G., Moreno, P.M., 2014. AquaCrop: FAO's crop water productivity and yield response model. Environ. Model. Software 62, 351–360.
Vapnik, V., Golowich, S., Smola, A., 1996. Support vector method for function approximation, regression estimation and signal processing. Adv. Neural Inf. Process. Syst. 9.
Varella, H., Guérif, M., Buis, S., 2010. Global sensitivity analysis measures the quality of parameter estimation: the case of soil parameters and a crop model. Environ. Model. Software 25, 310–319.
Visconti, F., de Paz, J., Martínez, D., Molina, M., 2014. Laboratory and field assessment of the capacitance sensors Decagon 10HS and 5TE for estimating the water content of irrigated soils. Agric. Water Manag. 132, 111–119. https://doi.org/10.1016/j.agwat.2013.10.005.
Wang, Xiaowen, Li, L., Ding, Y., Xu, J., Wang, Y., Zhu, Y., Wang, Xiaoyun, Cai, H., 2021. Adaptation of winter wheat varieties and irrigation patterns under future climate change conditions in Northern China. Agric. Water Manag. 243, 106409.
Witten, I.H., Frank, E., Hall, M.A., 2011. Data mining: practical machine learning tools and techniques. In: Morgan Kaufmann Series in Data Management Systems, third ed. Morgan Kaufmann, Burlington, MA.
Yan, W., Liu, W., Cheng, Z., Kan, J., 2010. The prediction of soil moisture based on rough set-neural network model. In: Presented at the Proceedings of the 29th Chinese Control Conference. IEEE, pp. 2413–2415.
Yin, X., Stam, P., Kropff, M., Schapendonk, H., 2003. Crop modeling, QTL mapping, and their complementary role in plant breeding. Agron. J. 95, 90–98. https://doi.org/10.2134/agronj2003.9000a.
Young, M.H., Fleming, J.B., Wierenga, P.J., Warrick, A.W., 1997. Rapid laboratory calibration of time domain reflectometry using upward infiltration. Soil Sci. Soc. Am. J. 61, 707–712. https://doi.org/10.2136/sssaj1997.03615995006100030001x.
Yu, Z., Liu, D., Lü, H., Fu, X., Xiang, L., Zhu, Y., 2012. A multi-layer soil moisture data assimilation using support vector machines and ensemble particle filter. J. Hydrol. 475, 53–64.
Zhang, B., Liu, Y., Xu, D., Zhao, N., Lei, B., Rosa, R.D., Paredes, P., Paço, T.A., Pereira, L. S., 2013. The dual crop coefficient approach to estimate and partitioning evapotranspiration of the winter wheat–summer maize crop sequence in North China Plain. Irrigat. Sci. 31, 1303–1316. https://doi.org/10.1007/s00271-013-0405-1.

Zhang, C., Xie, Z., Wang, Q., Tang, M., Feng, S., Cai, H., 2022. AquaCrop modeling to explore optimal irrigation of winter wheat for improving grain yield and water productivity. Agric. Water Manag. 266, 107580 https://doi.org/10.1016/j.agwat.2022.107580.

Zhang, H., Wang, S., Liu, K., Li, X., Li, Z., Zhang, X., Liu, B., 2022. Downscaling of AMSR-E soil moisture over north China using random forest regression. ISPRS Int. J. Geo-Inf. 11, 101.

Zheng, H., Shao, R., Xue, Y., Ying, H., Yin, Y., Cui, Z., Yang, Q., 2020. Water productivity of irrigated maize production systems in Northern China: a meta-analysis. Agric. Water Manag. 234, 106119.

Zhou, L., Liao, S., Wang, Z., Wang, P., Zhang, Y., Yan, H., Gao, Z., Shen, S., Liang, X., Wang, J., Zhou, S., 2018. A simulation of winter wheat crop responses to irrigation management using CERES-Wheat model in the North China Plain. J. Integr. Agric. 17, 1181–1193.

Zotarelli, L., Dukes, M., Scholberg, J., Femminella, K., Muñoz-Carpena, R., 2011. Irrigation scheduling for green bell peppers using capacitance soil moisture sensors. J. Irrigat. Drain. Eng. 137 https://doi.org/10.1061/(ASCE)IR.1943-4774.0000281.