Agricultural Water Management 328 (2026) 110300

Contents lists available at ScienceDirect

Agricultural Water Management

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

Uncertainty-aware irrigation scheduling based on probabilistic site-specific soil moisture predictions with SWIM^2: A case study in Flanders

Marit G.A. Hendrickx [a,b,*], Pieter Janssens [a,c,d], Jan Vanderborght [a,e], Evi Matthyssen [f], Anne Waverijn [g], Sander Bombeke [h], Jan Diels [a,b]

[a] Department of Earth and Environmental Sciences, KU Leuven, Leuven, 3001, Belgium
[b] KU Leuven Plant Institute (LPI), KU Leuven, Leuven, 3001, Belgium
[c] Soil Service of Belgium, Leuven, 3001, Belgium
[d] Department of Biosystems, KU Leuven, Leuven, 3001, Belgium
[e] Agrosphere Institute IBG-3, Forschungszentrum Jülich GmbH, Jülich, 52425, Germany
[f] Praktijkpunt Landbouw Vlaams-Brabant, Herent, 3020, Belgium
[g] Viaverda vzw, Kruishoutem, 9770, Belgium
[h] Proefstation voor de Groenteteelt, Sint-Katelijne-Waver, 2860, Belgium

ARTICLE INFO

Handling Editor - Enrique Fernández

Keywords:
Real-time irrigation scheduling
Probabilistic soil moisture forecasting
Bayesian inverse modeling
Probabilistic decision support

ABSTRACT

This study presents and evaluates a real-time decision support system (DSS) for site-specific irrigation scheduling based on soil moisture forecasting with SWIM^2 (Sensor Wielded Inverse Modeling of a Soil Water Irrigation Model). The SWIM^2 framework integrates a soil water balance model with in situ sensor data and soil moisture samples through Bayesian inverse modeling to generate probabilistic 10-day soil moisture forecasts. We assess the performance of the soil moisture forecasts and the irrigation DSS by integrating the model parameter ensemble with either deterministic or ensemble-based probabilistic weather forecasts, providing insights into their benefits and trade-offs in real-time irrigation management. Both approaches resulted in high detection rate and accuracy in predicting water stress triggering the irrigation threshold. The full ensemble yielded slightly better reliability at longer lead times whereas the probability distribution of the soil moisture predictions at short lead times was dominated by the SWIM^2 parameter uncertainty. Simulation of different irrigation treatments using the calibrated SWIM^2-based model illustrated and confirmed its potential for evaluating water use efficiency and crop response. Overall, this work illustrates the application and practical advantages of a probabilistic, ensemble-based modeling framework in supporting site-specific, data-informed irrigation strategies.

1. Introduction

Over the past 50 years, global irrigated areas have expanded by 70%, doubling water consumption, often in already water-stressed regions (Mehta et al., 2024). In Belgium, agricultural land covers 45% of the country, with only 1.8% designated as irrigable, and less than half of that actually irrigated (STATBEL, 2023; Eurostat, 2019). Increasing occurrences and prolonged periods of drought during the growing season over the past decade result in a higher crop water requirement. For high-value crops such as vegetables, supplementary irrigation is often required to maintain high yield and quality. Waverijn et al. (2024) demonstrated that, without irrigation, leek crops experienced significant drought stress in 2020–2021 in Flanders.

Global availability of commercial data-driven and software-based irrigation decision support systems (DSSs) for agricultural and horticultural crops is increasing (Zinkernagel et al., 2020). The integration of Internet-of-Things (IoT) and wireless sensor networks (WSNs) has played a significant role in advancing irrigation scheduling by enabling real-time soil moisture monitoring and automated decision-making (Sui, 2018; Alves et al., 2019; Jabro et al., 2020; El-Naggar et al., 2020). WSNs, often combined with machine learning techniques, have been used to forecast soil moisture and water stress, optimizing crop water use efficiency (Brinkhoff et al., 2019; Goap et al., 2018; Torres-Sanchez et al., 2020; Wang et al., 2025). However, machine learning approaches require large datasets, they struggle to generalize across unmonitored locations, and they often do not consider measurement errors.

* Correspondence to: Department of Earth and Environmental Sciences, KU Leuven, Celestijnenlaan 200e - bus 2411, Leuven 3001, Belgium.
E-mail address: marit.hendrickx@kuleuven.be (M.G.A. Hendrickx).

https://doi.org/10.1016/j.agwat.2026.110300
Received 28 October 2025; Received in revised form 12 February 2026; Accepted 15 March 2026
Available online 18 March 2026

Soil moisture models can simulate and predict soil moisture when observations are not available, providing us with short-term and seasonal water status predictions that can be integrated into DSSs to optimize irrigation management (Zinkernagel et al., 2020). Soil water balance (SWB) "bucket" models, like AquaCrop (Raes et al., 2009; Steduto et al., 2009) and BUDGET (Raes, 2002), offer a simplified approach that is easy to parameterize, making them ideal for real-time irrigation decisions with limited data. In contrast, Richards' equation-based models, such as HYDRUS (Simunek et al., 2005) and SWAP (Van Dam et al., 1997), provide more detailed simulations of soil water fluxes but require extensive input data and computational resources.

Soil moisture models can be parameterized in multiple ways, either through direct measurements or through calibration based on indirect measurements. For example, Vories and Sudduth (2021) explored sensor-based methods for determining field capacity in irrigation planning and suggested that apparent electrical conductivity measurements could serve as a viable alternative to extensive soil sampling. Bayesian methods have shown great potential for parameter estimation in soil hydrological modeling (Yang et al., 2007, 2015; Bulygina and Gupta, 2009; Schübl et al., 2022). Wang et al. (2021) proposed the Bayesian inversion of soil hydraulic properties from simplified evaporation experiments by fitting the model to observed pressure head and evaporation rate.

Most existing models, parameterizations, and decision support systems (DSSs) are deterministic. In contrast, the SWIM^2 framework (Sensor Wielded Inverse Modeling of a Soil Water Irrigation Model) produces probabilistic soil moisture predictions by integrating soil samples, real-time sensor data and their uncertainty into a soil water balance (SWB) model through Bayesian inverse modeling (Hendrickx et al., 2026). On top of this, weather forecasts can be integrated into the soil moisture predictions in real time. However, the uncertainty of weather forecasts, particularly for medium-range and localized predictions, can substantially affect soil moisture forecasts. Combining weather forecast uncertainty with parameter uncertainty in soil moisture models is crucial to ensure accurate estimates of prediction uncertainty. Error propagation further amplifies uncertainty in soil moisture forecasts, particularly for medium-range predictions of 7 days or more. Translating these probabilistic forecasts into actionable irrigation advice remains a key challenge. The final prediction uncertainty estimates provide more reliable guidance and enable risk-based decision-making for irrigation scheduling, accounting for prediction uncertainty, varying water availability and the diverse attitudes of farmers, ranging from conservative to risk-tolerant.

The objective of this study is to evaluate real-time soil moisture forecasting and irrigation scheduling using the SWIM^2 framework. The framework propagates uncertainty from in situ sensor and sample observations into model parameter distributions, resulting in probabilistic soil moisture estimates. We develop a real-time DSS that accounts for both observation and weather forecast uncertainty to support more robust irrigation decisions. Specifically, we quantify the uncertainty of 10-day soil moisture forecasts using deterministic and probabilistic weather inputs, and assess how this uncertainty influences irrigation scheduling. To illustrate its application, we simulate multiple irrigation treatments based on a model calibrated on a reference irrigated treatment and evaluate their effects on crop water use efficiency (WUE_c), irrigation efficiency (IE), and crop water productivity (WP_c). This study demonstrates how SWIM^2 can operate in a near-real-time setting to improve irrigation scheduling and water management in agricultural systems.

2. Materials and methods

2.1. Study sites

2.1.1. Case study of Belgian endive in Herent, 2023

An irrigation experiment was performed on a field of Belgian endive at *Praktijkpunt Landbouw Vlaams-Brabant* in Herent, Belgium (50°54′55"N, 4°41′44"E), during the growing season of 2023 (Fig. 1a). The experimental field covered 60 m by 15 m. Belgian endive was sown on June 5, 2023. This growth cycle was relatively wet, with a total precipitation of 310 mm over a period of 119 days. After the seed emergence stage, during which 81 mm of irrigation was applied, three different irrigation treatments were implemented using subsurface drip irrigation: a 100% irrigated treatment (T_{100}), receiving an additional 47 mm, a 50% irrigated treatment (T_{50}), receiving an additional 23.5 mm, and a zero-irrigation treatment (T_0), which received no additional irrigation. Each treatment was repeated in five plots of 3 m by 12 m, arranged in a fixed spatial pattern with treatments assigned per row (Fig. 1b). The field had a silt soil, with soil moisture sensors (see Section 2.2.1) installed in one of the T_{100} plots and soil moisture samples collected from multiple plots in each treatment. At the end of the growing season, all plots were harvested separately. The fresh weight and the number of tubers per plot were measured, which were used to calculate the tuber yield per unit area.

2.1.2. Case study of leek in Kruisem, 2022

Another irrigation experiment was performed on a field of leek at *Viaverda* in Kruisem, Belgium (50°56′42"N, 3°31′27"E), during the growing season of 2022 (Fig. 1a). The experimental field covered 81 m by 8.5 m. Leek was planted on May 5, 2022 and harvested on August 16, 2022. This growth cycle was much dryer compared to Herent in 2023, with only 127 mm of precipitation over a period of 203 days. Two irrigation treatments were implemented: a zero-irrigation treatment (T_0), and a 100% irrigated treatment (T_{100}), where 225 mm of irrigation was applied across 21 applications of subsurface drip irrigation throughout the season. Each treatment was repeated in four plots of 2.6 m by 8.5 m, arranged in a fixed spatial pattern with treatments assigned per row (Fig. 1c). The field had a loamy sand to sandy loam soil, with soil moisture sensors (see Section 2.2.1) installed in one of the T_{100} plots and soil moisture samples collected from multiple plots in each treatment. At the end of the growing season, all plots were harvested separately and the fresh yield per unit area was measured.

2.2. Measurements

2.2.1. Soil moisture sensor setup

The sensor measurement setup was identical to that described by Hendrickx et al. (2025, 2026). Dielectric capacitance soil moisture sensors (TEROS 10, METER Group, Inc., USA) were employed to measure volumetric soil water content (SWC) in the fields. These sensors use an electromagnetic field to assess the dielectric permittivity of the surrounding medium within a measurement volume of 430 mL, approximately corresponding to a cylinder with a diameter of 7.1 cm and a height of 10.9 cm. The raw sensor output S_{mV} (mV) was converted to soil moisture θ_{sensor} ($\mathrm{m^3 m^{-3}}$) using the manufacturer's calibration equation (METER Group, 2018):

$$\theta_{\mathrm{sensor}} = -2.154 + 3.898 \times 10^{-3} \times S_{\mathrm{mV}} - 2.278 \times 10^{-6} \times S_{\mathrm{mV}}^2 + 4.824 \times 10^{-10} \times S_{\mathrm{mV}}^3 \tag{1}$$

A single sensor module was installed in a T_{100} field plot (AgriSense Pro, Io-Things, Belgium). The module consisted of three TEROS 10 sensors connected to a datalogger with a Sigfox communication unit, which transmitted measurements to an online server for real-time access. The three sensors were installed horizontally at a depth of 15 cm within the crop row, spaced 2 m apart along the row. Sensor measurements were temporally resampled to daily observations by selecting the last measurement of each day.

To address discrepancies between point measurements from sensors installed at 15 cm depth and soil samples representing the entire 0 – 30 cm soil layer, the pooled sensor calibration from Hendrickx et al. (2025) was applied. This calibration curve was developed using an orthogonal fit of soil moisture sample measurements from the 0 – 30 cm

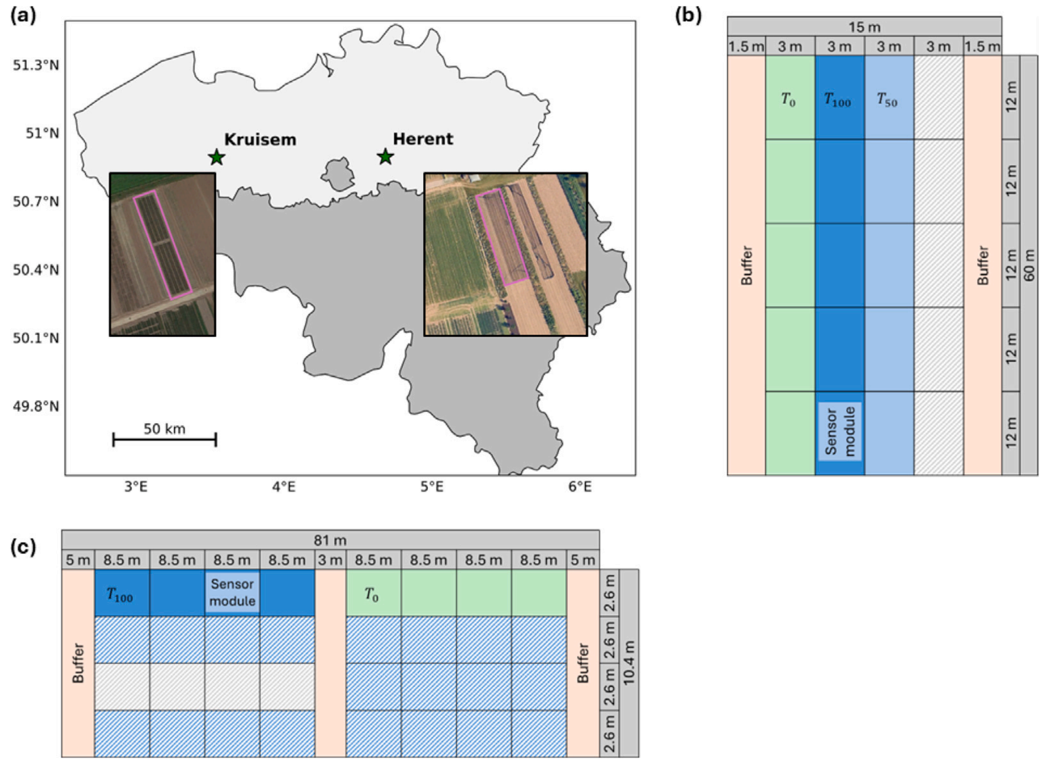

Fig. 1. Study sites are indicated on the map of Belgium (a). Plot distribution maps are shown for the Herent (b) and Kruisem (c) case study. The T_0, T_{50}, and T_{100} treatments are indicated, including the soil moisture sensor module installed in one T_{100} plot. The hatched area denotes a treatment not included in the present study. Blue-hatched areas received the same irrigation amount as T_{100} but were irrigated using different methods.

layer to their corresponding sensor measurements at 15 cm depth, based on data collected across multiple study sites in 2021, 2022, and 2023:

$$\theta_{0-30\ \mathrm{cm}} = -0.006\ + \bar{\theta}_{\mathrm{sensor}} \times 1.26 \tag{2}$$

where $\theta_{0-30\ \mathrm{cm}}$ ($\mathrm{m^3m^{-3}}$) represents the volumetric SWC in the $0-30$ cm soil layer, and $\bar{\theta}_{\mathrm{sensor}}$ ($\mathrm{m^3m^{-3}}$) is the average volumetric SWC measured by the three sensors at 15 cm depth. The calibration curve had an R^2 of 0.67 and an RMSE of 0.043 $\mathrm{m^3m^{-3}}$ (Hendrickx et al., 2025).

2.2.2. Soil samples

To characterize the soil properties, three undisturbed Kopecky ring samples (V: 100 $\mathrm{cm^3}$, h: 51 mm) were collected at each study site. Soil water retention was measured at various matric potentials using the pressure plate method, which provided the saturation point at pF[1] 0, the theoretical field capacity at pF 2, the critical threshold at pF 2.7, and the permanent wilting point at pF 4.2. Finally, the ring samples were dried to determine dry bulk density (BD).

Disturbed soil moisture samples were collected manually throughout the growing season and served as reference SWC measurements. Sampling was conducted at intervals of two to four weeks, ensuring coverage of all major crop growth stages and a wide range of soil moisture states. At each sampling event, a composite soil moisture sample was obtained from each irrigation treatment. The composite sample consisted of six individual gouge auger soil samples ($0-30$ cm depth), spatially distributed across the plot to capture within-plot variability. SWC was determined using the gravimetric method, with drying to constant weight ensuring unbiased readings. The volumetric SWC (θ_{vol}) was then derived from the gravimetric SWC (θ_{grav}) and BD:

$$\theta_{\mathrm{vol}} = \theta_{\mathrm{grav}} \times \mathrm{BD} \tag{3}$$

[1] $\mathrm{pF} = \log_{10}(|h|)$, where h is the soil water pressure head in cm.

2.2.3. Weather data

The study sites were equipped with a rain gauge to have in situ precipitation data. Additionally, precipitation radar data were used to fill missing data and false zeros. Meteorological data for reference potential evapotranspiration ($\mathrm{ET_o}$) estimation, including temperature, wind speed, relative humidity, air pressure, sunshine, solar radiation, and cloudiness, were obtained via OpenKMI from the nearest suitable RMI stations: Beitem (46.6 km) for Kruisem (2022) and Sint-Katelijne-Waver (21 km) for Herent (2023). Station selection was based on proximity and data availability, particularly humidity observations. Given the flat and homogeneous agro-climatic conditions of the region, the selected stations were considered representative for $\mathrm{ET_o}$ estimation. Daily $\mathrm{ET_o}$ was calculated based on the FAO-56 Penman–Monteith method (Allen et al., 1998), using the PyETo python package.

2.3. Modeling framework: SWIM^2

SWIM^2 is an irrigation DSS that combines real-time sensor data and periodic soil samples with a soil water balance model based on the FAO guidelines. Using the $\mathrm{DREAM_{(ZS)}}$ Bayesian inverse modeling algorithm (Vrugt, 2016), SWIM^2 estimates soil and crop parameters, providing soil moisture predictions along with uncertainty estimates (Hendrickx et al., 2026).

2.3.1. Soil water balance model

The SWB model is a single-layered "bucket" model based on FAO-56 methods (Allen et al., 1998) that calculates the daily soil water balance for the growing root zone using weather, soil, and crop parameters. The inflows are effective precipitation ($P_{\mathrm{eff}} = P - \mathrm{RO}$) considering runoff (RO), effective irrigation (I_{eff}) and capillary rise (CR) from groundwater, while the outflows are actual evapotranspiration ($\mathrm{ET_a}$), including crop transpiration and soil evaporation, and deep percolation (DP) to deeper soil layers. It assumes a uniform SWC throughout the profile and does not consider lateral flow. The model uses a dual crop coefficient, suitable for real-time irrigation scheduling with frequent water applications. A detailed description of the model is provided by Hendrickx et al. (2026).

Twelve uncertain parameters of the SWB model were estimated using DREAM$_{(ZS)}$, while auxiliary parameters were obtained from measurements or literature. The uncertain parameters include seven crop parameters, including three growth stage-specific basal crop coefficients ($K_{cb,ini}$, $K_{cb,mid}$, $K_{cb,end}$), three growth stage lengths (L_{ini}, L_{dev}, L_{mid}) and the maximum rooting depth ($z_{r,max}$), and five soil parameters, including SWC at field capacity (θ_{FC}), saturated hydraulic conductivity ($\ln(K_{sat})$), curve number (CN), maximum groundwater table depth ($z_{GWT,max}$) and initial SWC (θ_{ini}).

A critical water stress threshold (θ_{crit}) is defined based on the readily available water (RAW) as a fraction of the total available water (TAW):

$$\theta_{crit} = \theta_{FC} - \left[p_{table} + 0.04\ \left(5 - \mathrm{ET_m}\right)\right] \times \left(\theta_{FC} - \theta_{WP}\right) \tag{4}$$

where p_{table} is a tabulated crop-specific soil water depletion fraction for non-stressed conditions and for $\mathrm{ET_m} \approx 5$ mm day^{-1}, which can be found in the FAO-56 Irrigation and Drainage Paper (Allen et al., 1998). When SWC drops below θ_{crit}, the crop experiences water stress and actual evapotranspiration ($\mathrm{ET_a}$) becomes lower than the maximum potential evapotranspiration ($\mathrm{ET_m}$).

The selected parameters represent those with the strongest influence on root zone soil moisture dynamics and crop water use. Uniform prior distributions were assigned to all estimated parameters within physically plausible ranges derived from literature values, site characteristics, and preliminary sensitivity analyses. Prior ranges were chosen sufficiently wide to account for site-to-site variability while avoiding non-physical parameter combinations. A complete description of the SWIM^2 model structure, parameter sensitivity, and inverse modeling strategy is provided in Hendrickx et al. (2026).

2.3.2. Inverse modeling with DREAM$_{(ZS)}$

DREAM$_{(ZS)}$ is a Bayesian inverse modeling algorithm used to estimate uncertain parameters of the SWB model. The algorithm, implemented in Python 3 by Eric Laloy, is based on the DREAM$_{(ZS)}$ Matlab code (version 1.5). A formal likelihood function for heteroscedastic autocorrelated error residuals was applied as described in Hendrickx et al. (2026) with the pooled error covariance matrix as described in Hendrickx et al. (2025).

2.4. Weather forecast-based SWC predictions

2.4.1. Weather forecast data and processing

ECMWF (European Centre for Medium-Range Weather Forecasts) provides medium-range weather forecasts, up to 15 days ahead, at a horizontal resolution of 9 km (Roberts et al., 2024). The atmospheric model ensemble forecasting system consists of 51 members ($\mathcal{W}$). It includes one deterministic control forecast ($\mathrm{w_{CF}}$) and 50 additional forecasts, each generated with slightly perturbed initial conditions and modified model physics. $\mathcal{W}$ provides a probabilistic representation of possible weather outcomes, offering insights into forecast uncertainty and the probability of specific events such as heavy rainfall. The deterministic, unperturbed control forecast ($\mathrm{w_{CF}}$) becomes available earlier in the day than the full ensemble ($\mathcal{W}$), and may thus be preferred in the context of real-time SWC prediction and irrigation scheduling.

Nine meteorological variables were downloaded from the Meteorological Archival and Retrieval System (MARS) database: maximum and minimum temperature at 2 m in the last 6 h, 10 m U-component and V-component of wind, 2 m temperature, 2 m dewpoint temperature, surface net short-wave (solar) radiation, surface net long-wave (thermal) radiation, and total precipitation, all by 6 h steps, from which daily values were derived. The meteorological variables were used to compute the daily reference potential evapotranspiration, $\mathrm{ET_o}$, based on the FAO-56 Penman–Monteith method (Allen et al., 1998), using the PyETo python package.

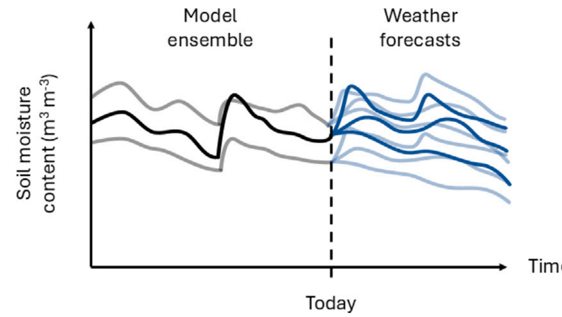

Fig. 2. Model parameter ensemble ($\mathcal{P}$) combined with the weather forecast ensemble ($\mathcal{W}$) resulting in $\Theta(\mathcal{W}, \mathcal{P})$.

A quantitative comparison between ECMWF forecasts and local meteorological observations for both $\mathrm{ET_o}$ and precipitation is provided in Appendix.

2.4.2. Integrating model parameter and weather forecast ensemble

Stratified random sampling was applied to the model parameter ensemble to reduce computational time when combining them with the 51 weather forecast members ($\mathcal{W}$). The post-burn-in iterations (T = 12 000) were divided into six sequential strata to mitigate the effects of autocorrelation in the Markov chains by ensuring that samples are drawn more evenly across the posterior distribution, rather than from dense, locally correlated regions. Within each stratum (2000), uniform random sampling selects $n = 200$ parameter sets (i.e., model parameter ensemble members), yielding a total of $N = 1200$ members in the model parameter ensemble ($\mathcal{P}$).

Multiple model prediction setups were constructed by combining the (subsampled) model parameter ensemble ($\mathcal{P}$) with various weather inputs to compare their impact on SWC forecasts and irrigation decisions (Table 1). Each of the model parameter ensemble members was run in combination with the weather control forecast ($\mathrm{w_{CF}}$) to get an ensemble of SWC predictions $\Theta(\mathrm{w_{CF}}, \mathcal{P})$. Each of the model parameter ensemble members was also run in combination with each $\mathcal{W}$ member ($N = 51$), resulting in a SWC prediction ensemble $\Theta(\mathcal{W}, \mathcal{P})$ of 61 200 members (Fig. 2). Finally, the maximum likelihood (ML) model parameter set was also run in combination with each $\mathcal{W}$ member (N = 51), resulting in a SWC prediction ensemble $\Theta(\mathcal{W}, \mathrm{p_{ML}})$ of 51 members.

Additionally, two reference ensembles were introduced. Each of the model parameter ensemble members ($\mathcal{P}$) was run in combination with the actually observed weather data as described in Section 2.2.3 during the forecasting period, resulting in a SWC prediction ensemble $\Theta(\mathrm{w_{OBS}}, \mathcal{P})$. A risk-averse scenario, $\Theta(\mathrm{w_{ZR}}, \mathcal{P})$, was introduced to represent conservative irrigation strategies, assuming zero rainfall during the 10-day forecast period and an evaporative demand ($\mathrm{ET_o}$) equal to the average of the past 10 days. This scenario is expected to predict dryer conditions compared to when applying $\mathcal{W}$ or $\mathrm{w_{CF}}$.

To assess the relative contributions of model parameter uncertainty and weather forecast uncertainty to SWC predictions, a total variance decomposition was performed. Ensembles of SWC predictions were generated by (i) combining the full model parameter ensemble ($\mathcal{P}$) with each individual weather forecast from $\mathcal{W}$, and (ii) combining each individual model parameter set with the full 51-member weather forecast ensemble ($\mathcal{W}$). The variance due to model parameter uncertainty was estimated from the spread of the model ensemble driven by a single weather forecast, while the variance due to weather forecast uncertainty was estimated from simulations using a single model parameter set combined with the full weather ensemble.

Table 1
Overview of ensembles with their abbreviation and number of members (N).

Abbreviation	Meaning	N
$\mathcal{W}$	Weather forecast ensemble	51
$\mathcal{P}$	Parameter ensemble	1200
$\Theta(\mathcal{W}, \mathcal{P})$	Prediction ensemble from parameter ensemble combined with weather forecast ensemble	61 200
$\Theta(\mathrm{w_{CF}}, \mathcal{P})$	Prediction ensemble from parameter ensemble combined with weather control forecast	1200
$\Theta(\mathrm{w_{OBS}}, \mathcal{P})$	Prediction ensemble from parameter ensemble combined with observed weather data	1200
$\Theta(\mathrm{w_{ZR}}, \mathcal{P})$	Prediction ensemble from parameter ensemble combined with zero rainfall weather scenario	1200
$\Theta(\mathcal{W}, \mathrm{p_{ML}})$	Prediction ensemble from ML parameter set combined with weather forecast ensemble	51

2.4.3. Real-time model calibration and SWC predictions

The model was calibrated at 5-day intervals (i.e., on day 20, 25, 30, ...) using all measurement data, both sensor and soil sample data, available up to that day, and SWC was predicted for the next 10 days by integrating the 10-day weather forecasts. For instance, on day 20, the calibration was performed using the past 20 daily averaged sensor measurements and a composite soil sample in that period. This 20-day period was defined as the calibration period. Then, SWC was predicted for the subsequent 10 days (e.g., days 21–30), which was defined as the prediction period. During the prediction periods, no irrigation was applied in the model, to simulate a decision-support context where future irrigation decisions are still to be made. However, the actual irrigation applications were included in the calibration periods, as part of the known system inputs. This process was repeated every 5 days until the end of the growth period, illustrating how the SWIM^2 framework can be applied in a near-real-time decision-support context. At each time step, the SWC prediction uncertainty and the probability of water stress were assessed.

2.5. Irrigation decision and simulation

2.5.1. Irrigation decision-making

To translate SWC predictions into actionable irrigation advice, a simple decision support system (DSS) was developed. The DSS uses the model-based estimates of water stress probability to support irrigation scheduling in the presence of uncertainty. The fraction of the SWC prediction ensemble members that surpass their critical water stress threshold can be interpreted as the probability of crop water stress (P_{stress}). This probability can be used as a risk indicator, with a cut-off (C) chosen based on factors such as water availability, irrigation costs or market conditions. This approach enables the exploration of probabilistic irrigation strategies that account for both model parameter uncertainty and weather forecast uncertainty.

In this study, we adopt a cut-off value of $C = 0.3$, meaning that irrigation is advised when at least 30% of the ensemble member predictions indicate the onset of water stress for at least two consecutive days. This value is used as an illustrative example of a moderately conservative, risk-averse irrigation strategy under conditions without explicit water availability constraints. Lower cut-off values would correspond to more risk-averse strategies that prioritize yield protection, while higher values would reflect more risk-tolerant strategies that prioritize water savings. The selected threshold is not intended to be universally optimal but demonstrates how probabilistic soil moisture forecasts can be translated into actionable irrigation advice within a flexible DSS framework.

The DSS is ideally run twice a week, for example on Mondays and Thursdays. In this study, P_{stress} was computed for the 10-day forecast period every five days.

2.5.2. Simulating irrigation treatments

Efficient irrigation management is crucial for optimizing water use in agriculture while maintaining crop productivity. In this study, we use the calibrated soil water balance model (SWIM^2) to simulate different irrigation treatments and evaluate their impact on crop water use efficiency ($\mathrm{WUE_c}$), irrigation efficiency (IE), and crop water productivity ($\mathrm{WP_c}$). The model is calibrated based on a reference treatment where a sensor module was installed. The parameter estimates are expected to be transferable to different irrigation treatments in the same field. First, the SWC simulations of the different irrigation treatments without in-situ sensors were evaluated using independent gravimetric soil moisture samples ($\mathrm{ME_{samp}}$ and $\mathrm{RMSE_{samp}}$). Then, the irrigation treatments can be evaluated based on simulated water use and yield results. By analyzing crop water use under varying irrigation strategies, we demonstrate the broader applicability of the modeling framework for improving the understanding of crop-water dynamics and informing irrigation decision-making beyond real-time operational use.

2.6. Evaluation metrics

2.6.1. Agricultural water productivity and efficiency metrics

Water efficiency metrics indicate how efficiently water input is utilized by the crop, or in other words, how effectively rainfall and irrigation are converted into actual evapotranspiration ($\mathrm{ET_a}$). Productivity metrics indicate how efficiently water is used by the crop to produce biomass, or in other words, how irrigation and actual evapotranspiration translate into biomass. Based on the SWIM^2 simulations, productivity and efficiency metrics can be calculated using observed rainfall (P), applied irrigation (I), and simulated $\mathrm{ET_a}$ in combination with observed yield (Y).

Crop water use efficiency ($\mathrm{WUE_c}$) is the ratio of effectively used water (i.e., (simulated) $\mathrm{ET_a}$) to the total water input (I and P), and indicates how efficiently the crop can take up the applied water and convert it into actual evapotranspiration:

$$\mathrm{WUE_c} = \frac{\mathrm{ET_a}}{I + P} \tag{5}$$

Crop water productivity ($\mathrm{WP_c}$) is the ratio of fresh crop yield (Y) to effective water use (i.e., (simulated) $\mathrm{ET_a}$), and indicates how efficiently the crop uses water:

$$\mathrm{WP_c} = \frac{Y}{\mathrm{ET_a}} \tag{6}$$

A higher $\mathrm{WP_c}$ means that the crop produces more biomass for the same amount of water use ($\mathrm{ET_a}$).

Irrigation water productivity ($\mathrm{WP_{irr}}$) is the ratio of additional crop yield (ΔY) to the additional amount of water applied through irrigation (ΔI), and provides an indication of the added value of the (additional) applied irrigation:

$$\mathrm{WP_{irr}} = \frac{\Delta Y}{\Delta I} \tag{7}$$

Irrigation efficiency (IE) is the ratio of beneficially used water (the additional $\mathrm{ET_a}$ resulting from irrigation) to the additional amount of water applied through irrigation (ΔI), and provides an indication of the added value of the applied irrigation:

$$\mathrm{IE} = \frac{\Delta \mathrm{ET_a}}{\Delta I} \tag{8}$$

We simulated the different treatments of the irrigation trials, which allowed us to calculate the difference in $\mathrm{ET_a}$ and compare it with the difference in irrigation amounts. An IE of 100% would mean that all of the (additional) irrigated water was used for (additional) evapotranspiration.

2.6.2. Prediction and classification accuracy of water stress and intervention

In contrast to deterministic forecasts, probabilistic forecasts are not simply "right" or "wrong". To assess the impact of probabilistic weather forecasts on SWC predictions for irrigation decisions, we assessed the predicted probability of water stress, and the binary classification of irrigation need, focusing on the model's ability to rank and discriminate between stressed and non-stressed conditions.

Discrimination skill was quantified using Receiver Operating Characteristic (ROC) curves, to evaluate the ability of the SWC forecasts to predict water stress probability (P_{stress}). The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1 – specificity) across a range of probability thresholds. A higher area under the curve (AUC) indicates better classification performance. An AUC of 1.0 corresponds to perfect discrimination between high and low probabilities of stressed conditions (P_{stress}), while an AUC of 0.5 indicates no skill (random classification). Note that the AUC metric only reflects the model's ability to correctly rank P_{stress}, regardless of the absolute values or a specific classification threshold. The reference probability of water stress, based on $\Theta(\mathrm{w}_{\text{OBS}}, \mathcal{P})$, was converted to a binary classification using a threshold of 0.3. While a threshold of 0.5 yields a balanced ROC curve, a lower threshold increases sensitivity to low-probability stress conditions, prioritizing the reduction of false negatives over false positives, and aligning with a more conservative irrigation strategy. In contrast, a higher threshold (e.g., 0.7) would focus the ROC analysis on the detection of high-probability or severe stress events. For each day in the forecast horizon, we compared the forecasted probability of water stress from the different model setups with the probability of water stress from $\Theta(\mathrm{w}_{\text{OBS}}, \mathcal{P})$. An ROC curve was computed for each forecast ensemble, i.e., $\Theta(\mathrm{w}_{\text{CF}}, \mathcal{P})$, $\Theta(\mathcal{W}, \mathcal{P})$, $\Theta(\mathcal{W}, \mathrm{p}_{\text{ML}})$ and $\Theta(\mathrm{w}_{\text{ZR}}, \mathcal{P})$, allowing us to compare how different model-weather input combinations affect the reliability of water stress probability detection. The curves were generated across all forecast days and averaged over the growing season to assess overall performance. This analysis highlights not only the predictive capability of the system but also its practical utility in real-time irrigation decision-making, where early and accurate detection of upcoming water stress is crucial for timely interventions.

In addition, we calculated standard classification metrics based on a fixed intervention threshold of 30% water stress probability to evaluate the accuracy of binary intervention predictions derived from the probabilistic model output. A prediction was classified as positive if the predicted water stress probability was greater than or equal to 0.3, representing a forecast of potential water stress conditions, while predictions below this threshold were classified as negative. To account for short-term forecasting uncertainty and to provide a more realistic evaluation of actionable predictions, ensuring these are not missed due to day-to-day variability, a 2-day sliding window approach was applied: intervention at a certain time point was considered positive if intervention was predicted on either the target day or the day after. The predicted binary classification was then compared against the binary reference based on $\Theta(\mathrm{w}_{\text{OBS}}, \mathcal{P})$.

To evaluate intervention classification performance, classification performance metrics were calculated, assessing the agreement between forecasted and reference water stress intervention while accounting for class imbalance and random agreement, where TP are True Positives, TN are True Negatives, FP are False Positives, and FN are False Negatives. Accuracy is defined as the proportion of correct predictions among all cases:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}\,, \tag{9}$$

while precision, or positive predictive value, is the proportion of predicted positives that are correct:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}\,, \tag{10}$$

and recall, sensitivity or true positive rate is the proportion of actual positives that are correctly identified:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}\,. \tag{11}$$

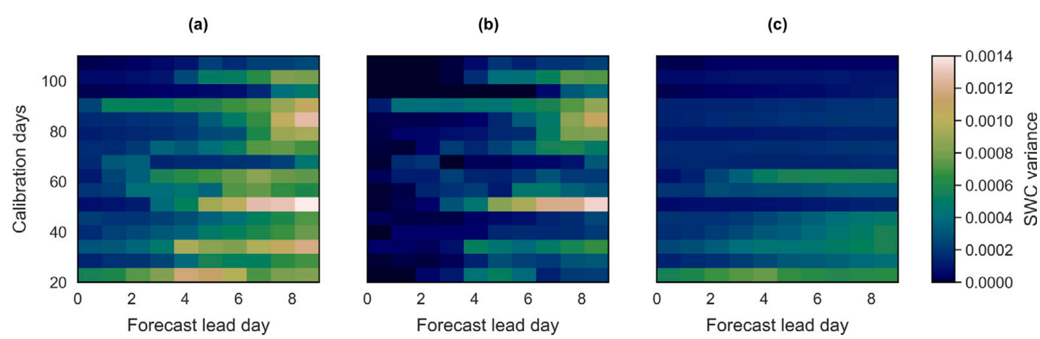

Fig. 3. Heatmaps of SWC variance (Herent, 2023) in function of forecast lead day and number of calibration days. (a) Total SWC variance, (b) SWC variance due to weather forecast uncertainty, and (c) SWC variance due to model parameter uncertainty.

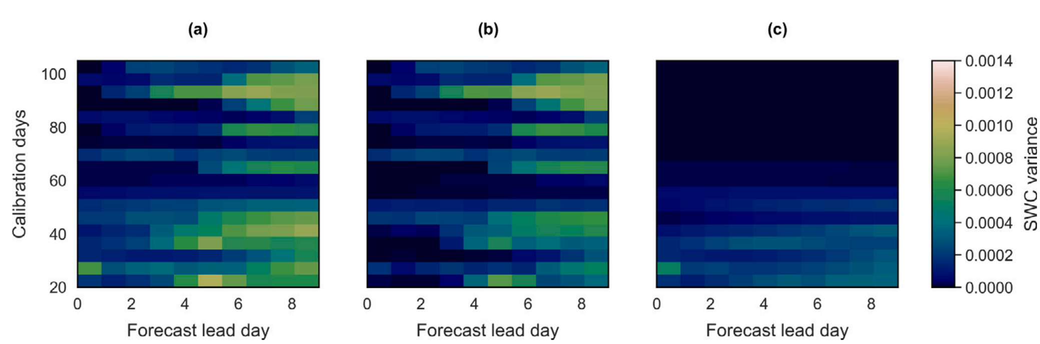

Fig. 4. Heatmaps of SWC variance (Kruisem, 2022) in function of forecast lead day and number of calibration days. (a) Total SWC variance, (b) SWC variance due to weather forecast uncertainty, and (c) SWC variance due to model parameter uncertainty.

Finally, the F1 score is the harmonic mean of precision and recall:

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}, \tag{12}$$

and Cohen's Kappa is a more complex measure of agreement between predicted and reference classifications, adjusted for agreement by chance.

Additionally, the root mean square error (RMSE) for the water stress probability forecast quantifies the discrepancy between the forecasted water stress probability ($\hat{P}_{\mathrm{stress},i}$) and the reference water stress probability ($P_{\mathrm{ref},i}$) at time i:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{P}_{\mathrm{stress},i} - P_{\mathrm{ref},i})^2} \tag{13}$$

where the reference water stress probability, $P_{\mathrm{ref},i}$, was based on $\Theta(\mathrm{w}_{\mathrm{OBS}}, \mathcal{P})$. This RMSE was computed for each lead time across all 10-day prediction periods, with N being the number of prediction periods. To quantify the uncertainty in the RMSE estimate, a 95% confidence interval was computed based on a leave-one-out (LOO) cross-validation approach, leaving out one prediction period each time.

3. Results

3.1. Impact of model parameter and weather forecast uncertainty on SWC prediction uncertainty

The contributions of weather forecast uncertainty and model parameter uncertainty to the SWC forecast variance were quantified using a total variance decomposition. Overall, the SWC variance due to parameter uncertainty slightly increases with lead time, but substantially decreases with an extending calibration period (Fig. 3c–4c), while SWC variance due to weather forecast uncertainty is independent of the calibration period (Fig. 3b–4b). As a result, the contribution of parameter uncertainty to the total SWC variance (Fig. 3a–4a) decreases as more days of SWC measurements are used in the SWIM^2 model calibration, and the contribution of the weather forecast uncertainty increases as a consequence. Moreover, as lead time increases, the weather forecast uncertainty and its contribution to SWC uncertainty increases (Fig. 3b–4b).

The sum of the SWC variance due to weather forecast uncertainty and model parameter uncertainty was, on average, comparable to the total SWC variance. Interaction terms were generally small, with negative values up to 0.005 $\mathrm{m^3 m^{-3}}$ (not shown). At the level of individual ensemble members, substantial interaction effects were observed, where uncertainties could amplify one another (positive interactions) or dampen each other (negative interactions).

Table 2
Confusion matrix components for water stress predictions with a threshold of 0.3 and applying a 2-day sliding window, for the two case studies combined, with TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives.

Model	TP	TN	FP	FN
$\Theta(\mathcal{W}, \mathcal{P})$	124	164	17	1
$\Theta(\mathrm{w}_{\mathrm{CF}}, \mathcal{P})$	121	164	17	4
$\Theta(\mathrm{w}_{\mathrm{ZR}}, \mathcal{P})$	123	143	38	2
$\Theta(\mathcal{W}, \mathrm{p}_{\mathrm{ML}})$	119	154	27	6

3.2. Uncertainty-based water stress prediction for irrigation decision-making

SWC prediction uncertainty strongly depends on weather forecast uncertainty as lead time increases (Section 3.1), and is expected to be different when using a single control weather forecast, w_{CF}, or the full weather ensemble, $\mathcal{W}$. The RMSE of predicted water stress probabilities was evaluated for $\Theta(\mathrm{w}_{\mathrm{CF}}, \mathcal{P})$, $\Theta(\mathcal{W}, \mathcal{P})$, and $\Theta(\mathcal{W}, \mathrm{p}_{\mathrm{ML}})$, as well as for the risk-averse scenario $\Theta(\mathrm{w}_{\mathrm{ZR}}, \mathcal{P})$ (Fig. 5). The 95% confidence interval on the RMSE, computed using a LOO approach over prediction periods, was largely positioned below the mean RMSE line. This is because most LOO RMSE values were similar, except for one LOO step where RMSE was substantially lower, likely due to the exclusion of a prediction period with much lower accuracy. As a result, the mean RMSE lies near the upper edge of the distribution, and the CI appears asymmetrically below it.

ROC curves for $\Theta(\mathcal{W}, \mathcal{P})$, $\Theta(\mathcal{W}, \mathrm{p}_{\mathrm{ML}})$, $\Theta(\mathrm{w}_{\mathrm{CF}}, \mathcal{P})$, and $\Theta(\mathrm{w}_{\mathrm{ZR}}, \mathcal{P})$ were computed by combining data from the two case studies and are shown in Fig. 6a. Among the evaluated approaches, $\Theta(\mathcal{W}, \mathcal{P})$ exhibited the highest AUC, along with the ROC curve bending sharply toward the top-left corner. The remaining prediction ensembles showed similar ROC curves with slightly lower AUC values. $\Theta(\mathrm{w}_{\mathrm{ZR}}, \mathcal{P})$ resulted in the lowest AUC.

Table 2 summarizes the confusion matrix for water stress predictions with a threshold of 0.3 using a 2-day sliding window for the two case studies combined. The full ensemble model, $\Theta(\mathcal{W}, \mathcal{P})$, achieved the highest overall performance, with an accuracy of 0.94, a precision of 0.88, and a recall of 0.99 (Fig. 6b). The model correctly identified nearly all water stress events while maintaining a relatively low false positive rate. The model ensemble using the deterministic control forecast, $\Theta(\mathrm{w}_{\mathrm{CF}}, \mathcal{P})$, showed slightly lower performance but still maintained high recall and precision. $\Theta(\mathcal{W}, \mathrm{p}_{\mathrm{ML}})$ showed a drop in all metrics, while the risk-averse model, $\Theta(\mathrm{w}_{\mathrm{ZR}}, \mathcal{P})$, achieved a high recall (0.98), detecting almost all stress cases that trigger a positive irrigation advice, but at the cost of reduced precision (0.76) due to a large number of false positives (Table 2).

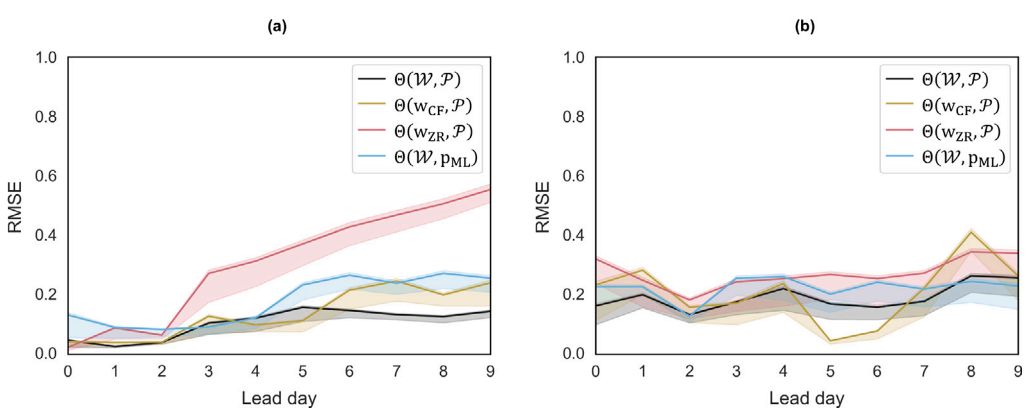

Fig. 5. Mean RMSE for the water stress probability forecast with a 95% CI for the case study in Herent, 2023 (a) and the case study in Kruisem, 2022 (b). The four colors indicate different prediction ensembles (Table 1).

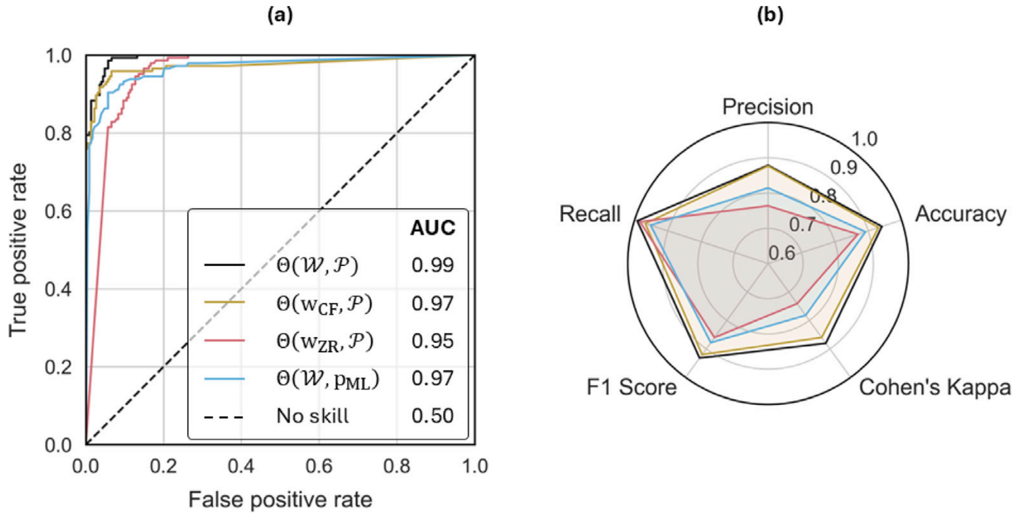

Fig. 6. (a) ROC curves for $\Theta(\mathcal{W},\mathcal{P})$, $\Theta(\mathcal{W},\mathrm{p_{ML}})$, $\Theta(\mathrm{w_{CF}},\mathcal{P})$, and $\Theta(\mathrm{w_{ZR}},\mathcal{P})$, across all lead days and for the two case studies combined, with their area under the curve (AUC). (b) Contingency metrics derived from Table 2, visualized as a radar plot.

3.3. Simulating irrigation treatments: Model-based evaluation of irrigation efficiency and crop productivity

3.3.1. Case study of Belgian endive in Herent, 2023

For the case study of Belgian endive in Flanders (2023), the model was calibrated based on one of the T_{100} plots, where a sensor module was installed (Fig. 7a). The T_{50} and T_0 treatment were simulated and are shown in Fig. 7b and Fig. 7c, respectively. Soil moisture simulations for irrigation treatments without in-situ sensors were evaluated using independent gravimetric soil moisture samples. Overall, simulated root zone soil moisture showed good agreement with observations, with deviations generally within the range of sampling uncertainty and an $\mathrm{RMSE_{samp}}$ of 0.018 $\mathrm{m^3m^{-3}}$ and 0.031 $\mathrm{m^3m^{-3}}$ for treatment T_{50} and T_0, respectively. However, more or continuous measurements would be needed to validate them more thoroughly.

Even though 2023 had some short periods of drought in Flanders, it was not a substantially dry growing season. This is why irrigation requirements were low and the zero-irrigation treatment still resulted in high yields and limited water stress (Table 3). However, a one-way ANOVA test and a Tukey HSD post-hoc test showed that the T_0 treatment had a significantly lower yield compared to T_{50} and T_{100}, indicating that irrigation did have a statistically significant effect on yield. The actual crop ET was simulated based on the ML model, and crop water stress was derived from this (Table 3). The actual-to-potential crop ET ratio increased with an increasing amount of additional irrigation, reaching near 100% of the water demand for treatment T_{100}.

3.3.2. Case study of leek in Kruisem, 2022

For the case study of leek in 2022, the model was calibrated based on one of the drip-irrigated T_{100} plots, where a sensor module was installed (Fig. 8a). Additionally, the T_0 treatment was simulated and can be compared to the samples (Fig. 8b). Note that the model simulation slightly under-predicts soil moisture for T_0 when compared to the soil moisture samples, with a $\mathrm{ME_{samp}}$ of 0.042 $\mathrm{m^3m^{-3}}$.

In contrast with 2023, the cumulative precipitation deficit during the growing season of 2022 in Flanders was substantial (221 mm over a period of 90 days) (VMM, 2025). The measured and model-derived results of the leek growth cycle in Kruisem (2022) are shown in Table 4. The 100% irrigated treatment had a significantly higher yield (+84%) compared to the non-irrigated treatment. The actual-to-potential crop ET ratio was 33% for the non-irrigated treatment, substantiating the occurred drought stress, and 97% for the irrigated treatment (T_{100}).

Table 3
Irrigation experiment in a field of Belgian endive at *Praktijkpunt Landbouw Vlaams-Brabant*, Herent in 2023. During the seed emergence stage, 81 mm irrigation was applied to all treatments. Different irrigation treatments were applied after the seed emergence phase: T_0 was the zero-irrigation treatment, T_{50} was the 50% irrigated treatment, and T_{100} was the 100% irrigated treatment. Crop water stress was computed for each treatment based on the ML model simulation of crop ET. At time of harvest, tuber yield results were gathered.

	T_0	T_{50}	T_{100}
Additional irrigation after emergence (mm)	0	23.5	47
Actual-to-potential crop ET ratio (%)†	95.6	98.2	99.9
Number of days with water stress†	18	11	3
Tuber yield (ton $\mathrm{ha^{-1}}$)	38.3	41.9	43.4
$\mathrm{WUE_c}$ (%)†	85.8	82.7	79.3
$\mathrm{WP_c}$ (ton $\mathrm{mm^{-1}\ ha^{-1}}$)†	1.1	1.2	1.2
$\mathrm{WP_{irr}}$ (ton $\mathrm{mm^{-1}\ ha^{-1}}$)	–	1.8	0.9
IE (%)†	–	36.1	29.3

† Model-based.

4. Discussion

4.1. Impact of model parameter and weather forecast uncertainty on SWC prediction uncertainty

Our results provide a quantitative basis for interpreting the sources of SWC prediction uncertainty, indicating that soil moisture prediction uncertainty is primarily driven by model parameter uncertainty at short lead times and by weather forecast uncertainty at longer lead times. The contribution of weather forecast uncertainty may depend on cumulative precipitation rather than lead time, since a large precipitation uncertainty may result in a large SWC uncertainty. Additionally, precipitation forecast uncertainties are larger compared to the uncertainties in evaporative demand forecasts, and thus contribute more to SWC uncertainty.

The total SWC variance was in general larger for the case study in Herent (2023) (Fig. 3a) compared to the case study in Kruisem (2022) (Fig. 4a). This is possibly due to the drier soil moisture conditions in 2022, and the low SWC variance due to parameter uncertainty under these conditions (Fig. 4c).

Furthermore, substantial interaction between parameter and weather forecast uncertainty among individual ensemble members indicates that certain model configurations are more sensitive to weather variations than others. For example, a larger soil water storage capacity

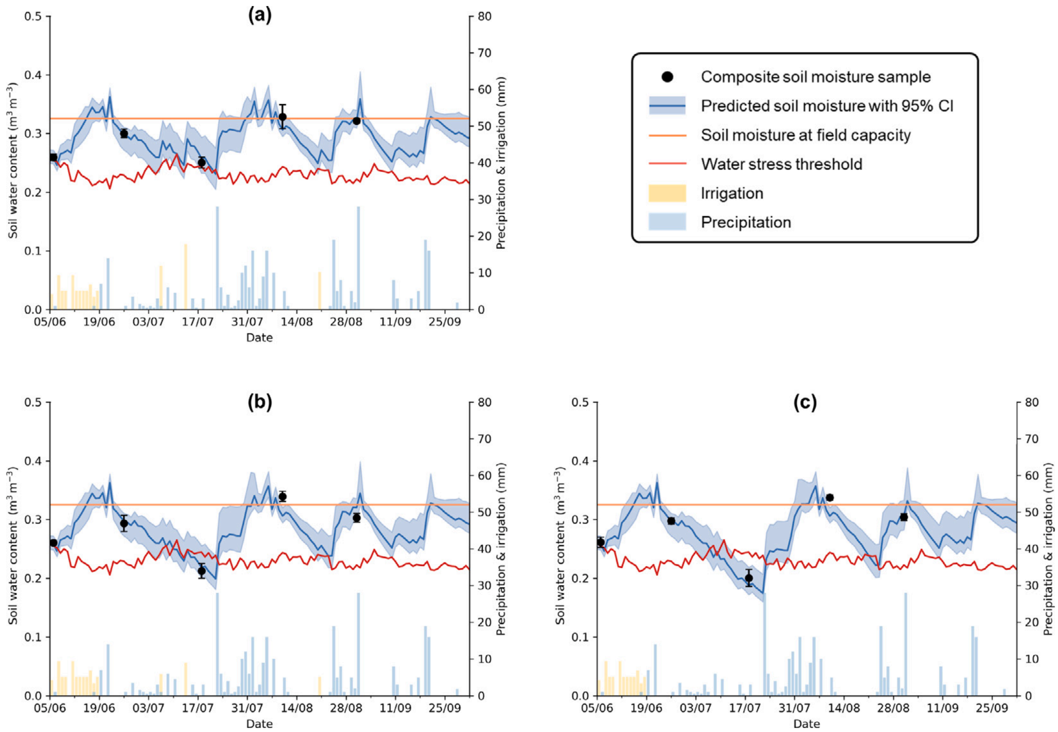

Fig. 7. Irrigation experiment in a field of Belgian endive at PLV, Herent in 2023. Different irrigation treatments were applied after the seed emergence phase (a: 100% irrigation, $RMSE_{samp}$ = 0.011 $m^3\,m^{-3}$; b: 50% irrigation, $RMSE_{samp}$ = 0.018 $m^3\,m^{-3}$; c: 0% irrigation, $RMSE_{samp}$ = 0.031 $m^3\,m^{-3}$).

due to a higher field capacity in combination with high precipitation could result in a positive interaction, while a low field capacity might dampen precipitation error effects. Since negative interactions were more prevalent, total SWC variance was slightly lower than the sum of variances.

Finally, it is important to note that SWC uncertainty does not follow a normal distribution, and instead often has a skewed distribution, or one with heavy tails. Heavy tails in prediction distributions mean more uncertainty or risk of extremes, which is important for risk modeling and decision-making, especially in agriculture and hydrology.

4.2. Integrating weather forecasts for uncertainty-based water stress prediction for irrigation decision-making

The operational applicability of forecast-driven irrigation support systems depends not only on accuracy, but also on computational and data-processing requirements. The weather control forecast only takes about 15 min to download from the MARS database for one location while the 50 perturbed ensemble members require about 10 h to download. Additionally, when the weather forecasts are combined with our model ensemble, using $\mathcal{W}$ requires 51 times more computational time compared to using w_{CF}. When parallelized over 6 processors, the computational time per prediction setup was approximately 1 min for the model parameter ensemble ($\mathcal{P}$) combined with a single weather forecast (e.g., w_{CF}), compared to approximately 51 min when integrated with the full weather forecast ensemble ($\mathcal{W}$), while the Bayesian estimation takes about 1 h for any individual case. In practice, w_{CF} is thus preferred in the context of real-time SWC prediction and irrigation scheduling, which is why we need to assess the trade-off in SWC prediction uncertainty and irrigation management decisions when using w_{CF} instead of $\mathcal{W}$.

Overall, RMSE of predicted water stress probabilities was lowest for $\Theta(\mathcal{W}, \mathcal{P})$, while RMSE for $\Theta(\mathcal{W}, p_{ML})$ was somewhat higher, indicating that the model parameter ensemble prediction outperformed the most likely scenario parameter set. This suggests that probabilistic decision-making can enhance predictive performance by better accounting for model parameter uncertainty. The RMSE was similar for $\Theta(w_{CF}, \mathcal{P})$ and $\Theta(\mathcal{W}, \mathcal{P})$ up to five lead days, while after 5 lead days, performance of $\Theta(w_{CF}, \mathcal{P})$ decreased or increased, possibly depending on weather conditions. The risk-averse scenario, $\Theta(w_{ZR}, \mathcal{P})$, showed the largest RMSE, especially during the wet growing season of 2023 (Fig. 5a), due to an underestimation of precipitation.

The higher AUC and pronounced curvature of the ROC curve toward the top-left corner indicate that $\Theta(\mathcal{W}, \mathcal{P})$ has a stronger ability to discriminate between stressed and non-stressed conditions compared to the other approaches. While the remaining ensembles also show acceptable discrimination skill, their slightly lower AUC values suggest reduced reliability. The lower performance of $\Theta(w_{ZR}, \mathcal{P})$ is primarily associated with an increased false positive rate, implying a tendency to predict water stress in non-stressed conditions. In an operational irrigation context, this could lead to unnecessary irrigation events. The wet conditions in 2023 led to mostly true negatives (TN), i.e., correct negative irrigation advice, while the drier conditions in 2022 resulted in more true positives (TP). The results demonstrate a trade-off between detecting all true water stress events (recall) and avoiding unnecessary irrigation triggers (precision). While the full ensemble model offers the best balance, the risk-averse scenario prioritizes detection of all stress events, which may lead to over-irrigation if applied operationally. This highlights the importance of considering both recall and precision when designing irrigation decision-support systems. The water stress threshold of 0.3 means that when a water stress probability of 30% is reached, irrigation advice is positive. In operational mode, the threshold can be adjusted depending on varying water availability and diverse attitudes of farmers, ranging from risk-averse to risk-tolerant.

In summary, these results clarify the practical trade-offs involved in choosing between different forecast types for real-time irrigation scheduling. Using the full weather ensemble in $\Theta(\mathcal{W}, \mathcal{P})$ provided the most accurate and reliable water stress predictions, while the control forecast model $\Theta(w_{CF}, \mathcal{P})$ achieved comparable performance with far lower computational demands. The risk-averse scenario $\Theta(w_{ZR}, \mathcal{P})$ successfully detected nearly all stress events but produced many false

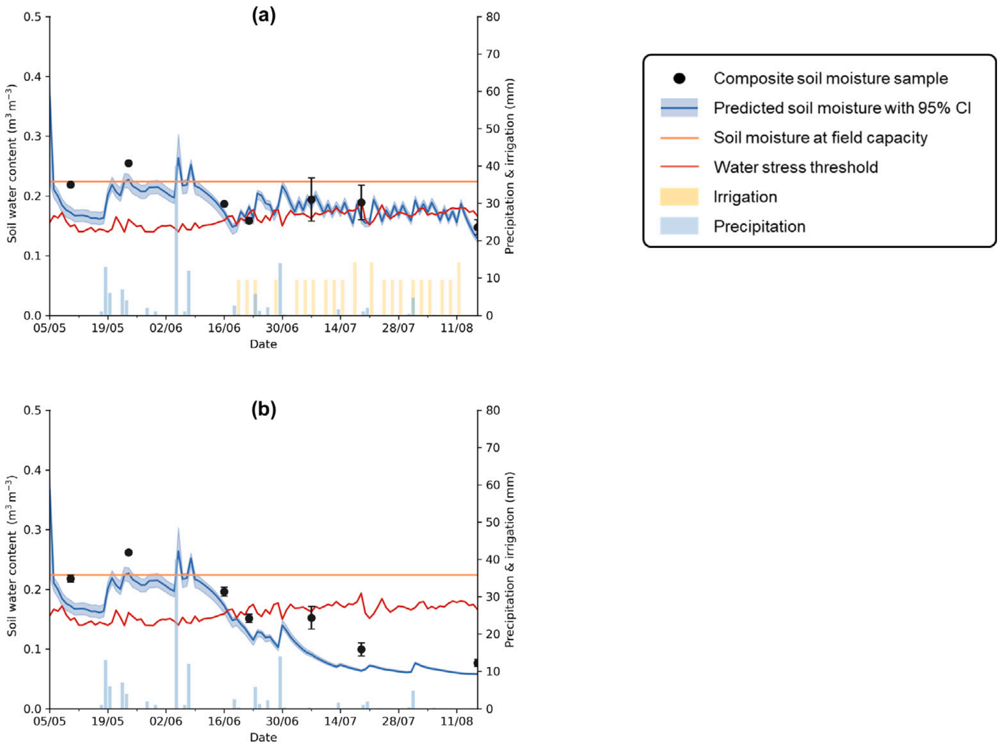

Fig. 8. Irrigation experiment in a field of leek at *Viaverda*, Kruisem in 2022. Different irrigation treatments were applied (a: 100% irrigation, $RMSE_{samp}$ = 0.032 $m^3\ m^{-3}$; b: 0% irrigation, $RMSE_{samp}$ = 0.045 $m^3\ m^{-3}$).

Table 4
Irrigation experiment in a field of leek at *Viaverda*, Kruisem in 2022. T_0 was the zero-irrigation treatment, and T_{100} was the 100% irrigated treatment. Crop water stress was computed for each treatment based on the ML model simulation of crop ET. At time of harvest, fresh weight yield results were gathered.

	T_0	T_{100}
Additional irrigation after emergence (mm)	0	225
Actual-to-potential crop ET ratio (%)[†]	33.0	96.7
Number of days with water stress[†]	59	23
Yield (ton ha^{-1})	20.3	41.1
WUE_c (%)[†]	113.3	95.6
WP_c (ton mm^{-1} ha^{-1})[†]	1.4	1.2
WP_{irr} (ton mm^{-1} ha^{-1})	–	0.9
IE (%)[†]	–	90.1

[†] Model-based.

positives, leading to potential over-irrigation. These findings demonstrate that ensemble-based forecasts offer the highest predictive value where feasible, while w_{CF}-based strategies provide an efficient near-optimal alternative for operational use. The profit that is linked to these setups however depends on the cost of an irrigation event, and more specifically the cost of false positives (irrigated when not needed) versus false negatives (no irrigation when irrigation was needed).

4.3. Simulating irrigation treatments: Model-based evaluation of irrigation efficiency and crop productivity

In Herent 2023, simulated SWC showed good agreement with the soil moisture samples across irrigation treatments. This indicates that the parameter sets calibrated on the T_{100} treatment provide a reasonable representation of soil water dynamics under different irrigation regimes. The transfer of calibrated parameters from the T_{100} treatment to other irrigation treatments assumes that soil-related parameters, such as field capacity (θ_{FC}) and saturated hydraulic conductivity (K_{sat}), remain invariant across plots, which is justified given the homogeneous soil texture within each experimental site. In contrast, crop-related parameters, particularly basal crop coefficients and growth stage development, may vary between treatments due to water stress effects and differences in canopy development. This effect is illustrated in Fig. 8 for leek in 2022, where the non-irrigated treatment exhibited systematically higher observed soil moisture than simulated values. The reduced yield of this treatment, approximately half that of the irrigated plots, suggests limited canopy development and a lower crop coefficient. As a result, ET_a may have been overestimated, leading to an underestimation of simulated soil water content relative to the soil moisture samples.

In the context of irrigation and crop yield, the law of diminishing returns implies that, as irrigation increases, the yield initially increases sharply, but beyond a certain point, additional irrigation results in smaller and smaller yield gains, and eventually may even reduce yield due to waterlogging, nutrient leaching, or disease. This is reflected in the observed decline of WUE_c and IE (%) in Herent as irrigation amounts increase, since the additional water applied between T_{50} and T_{100} produced a smaller yield gain compared to the same amount of water applied between T_0 and T_{50}. In Kruisem, the high IE of 90% means that almost all applied irrigation water in T_{100} contributed to additional evapotranspiration. This indicates that irrigation was primarily applied when soil moisture was limiting (i.e., during stress), with little to no water "lost" due to application during non-stress periods.

A direct comparison of WUE and IE for Belgian endive was not possible, as, to our knowledge, no literature values are available. Cheng et al. (2021) reported larger yield losses but higher WUE under deficit irrigation for vegetables compared with field crops, providing context for the relatively high irrigation demand observed in Belgian endive and leek production. Waverijn et al. (2024) obtained a WP_{irr} of 0.2 ton mm^{-1} ha^{-1} for leek, based on an irrigation amount of 75 mm. This value is substantially lower than the WP_{irr} of 0.9 ton mm^{-1} ha^{-1} obtained in the present study. The lower irrigation water productivity of Waverijn et al. (2024) can be explained by the limited yield response to irrigation, as their non-irrigated treatment already exhibited a high actual-to-potential crop ET ratio (94%), indicating only minor water stress.

Although on-farm performance gains of SWIM^2 relative to traditional irrigation were not quantified here, the framework illustrates a pathway toward reducing over-irrigation and decision-making burden by explicitly accounting for soil moisture dynamics and forecast uncertainty. Additionally, because the framework is driven by soil moisture observations from a single measurement zone of 80 m^2, its applicability to large, heterogeneous fields is limited; addressing this will require spatial monitoring or extrapolation approaches.

4.4. Future perspectives and application

Future research may include expansion of the framework to larger spatial scales, integrating remotely sensed data for spatial soil moisture information, and exploring adaptive decision rules that dynamically adjust irrigation thresholds based on evolving forecast uncertainty and crop development stages. Additionally, the SWIM^2 framework could benefit from including crop yield predictions using a similar approach as the AquaCrop model (Raes et al., 2009; Steduto et al., 2009) by integrating a yield module where biomass is simulated based on crop water productivity and crop transpiration, and final yield depends on a stress-dependent harvest index. In-season yield forecasting could be based on seasonal soil moisture predictions based on extended-range or seasonal weather forecasts, historical weather data, or climatology (Boas et al., 2023; Newlands et al., 2014; Togliatti et al., 2017). Since SWIM^2 currently has low information gain on crop growth dynamics when using only SWC observations, additional data sources such as canopy cover or vegetation indices (e.g., NDVI) derived from optical remote sensing are recommended. These data could be used to dynamically constrain crop growth parameters and treatment-specific crop coefficients, allowing improved representation of canopy development and water stress effects across contrasting irrigation regimes. Such integration would enhance the simulation of crop transpiration, biomass accumulation, and yield.

Beyond methodological extensions, future work should also address the broader applicability and operational scalability of the framework across regions with varying data availability and user capacities. The proposed approach is designed with operational use in mind: from the farmer's perspective, interaction with the system is limited to a simple user interface that displays real-time sensor observations, short-term SWC predictions, and a clear irrigation recommendation. The underlying modeling, uncertainty propagation, and probabilistic analysis are fully automated and do not require technical expertise from the end user. In the present study, the system is embedded within the Soil Service of Belgium, which provides data quality control, system maintenance, and user support, thereby lowering the barrier for adoption by farmers. Regarding weather data accessibility, the framework is flexible and can be adapted to different spatial scales and regions. In Belgium and Flanders, high-resolution precipitation radar data and a dense network of weather stations are available nationwide, and the relatively homogeneous climate allows meteorological data to be considered representative over moderate spatial distances. At the European scale, regions lacking dense in situ observations can rely on ECMWF reanalysis products, although local validation and calibration would be required. To further enhance scalability, ongoing and future research will focus on developing a European-level irrigation DSS framework in which SWIM^2 serves as a core modeling component, facilitating harmonized implementation across regions with contrasting climatic conditions and data availability.

5. Conclusions

This study introduced a probabilistic framework for real-time soil moisture forecasting and irrigation scheduling using the SWIM^2 model. The results showed that uncertainty in soil moisture predictions arises mainly from model parameters at short lead times and from weather forecasts at longer lead times. Ensemble-based forecasts provided the most accurate and reliable predictions of water stress, while control forecast strategies offered a practical balance between predictive performance and computational cost. The risk-averse approach captured nearly all stress events but often triggered unnecessary irrigation, emphasizing the importance of balancing reliability and efficiency in operational decision-making. Simulated irrigation treatments confirmed the framework's ability to evaluate water use efficiency and crop response under different conditions. Overall, the findings highlight the potential of ensemble-based, site-specific soil moisture modeling to support more informed and adaptive irrigation scheduling in precision agriculture.

CRediT authorship contribution statement

Marit G.A. Hendrickx: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Pieter Janssens:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. **Jan Vanderborght:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Evi Matthyssen:** Resources, Data curation. **Anne Waverijn:** Resources, Data curation. **Sander Bombeke:** Resources, Data curation. **Jan Diels:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author(s) used ChatGPT (GPT-5, OpenAI) in order to assist with phrasing and improve the readability of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Marit Hendrickx reports financial support was provided by Research Foundation Flanders. Pieter Janssens reports financial support was provided by Flanders Innovation & Entrepreneurship. 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.

Acknowledgments

This work was funded by the Research Foundation Flanders, Belgium (FWO fellowship 1S20822N). The study sites were monitored in the context of a project funded by Flanders innovation & entrepreneurship, Belgium (VLAIO project HBC.2018.2201). We acknowledge the support of the Soil Service of Belgium, who collected the field data, as well as research centres in Flanders as partners (*Viaverda*, *Proefstation voor de Groenteteelt* and *Praktijkpunt Landbouw Vlaams-Brabant*). Additionally, we want to thank Christophe Coeck (KU Leuven) for the technical and practical support concerning the sensor module and IoT setup, and Lore Fondu (KU Leuven) for the support with soil analyses in the lab.

Appendix. Comparison of weather data and ECMWF forecasts

A comparison was conducted between reference evapotranspiration (ETo) calculated from RMI station observations and ETo derived from ECMWF ensemble forecasts. For Herent (2023), station-based ETo values were generally lower than the corresponding ECMWF forecasts, with a positive forecast bias of approximately 0.5–0.9 mm. Observed ETo values fell within the lower 10th percentile of the forecast ensemble distribution for short lead times and within the lower 25th percentile for longer lead times. For Kruisem (2022), good agreement was observed between station-based and ECMWF-derived ETo. No systematic bias was detected, and station-based ETo values were typically located between the 40th and 60th percentiles of the ensemble distribution, depending on lead time. ECMWF forecasts have a spatial resolution of approximately 9 km, whereas the RMI stations were located at Beitem (46.6 km from Kruisem) and Sint-Katelijne-Waver (21 km from Herent). Differences between station-based observations and gridded forecasts may therefore partly reflect spatial variability in meteorological conditions.

Additionally, precipitation from rain gauge measurements, complemented with radar-based observations, was compared with ECMWF ensemble forecasts. Five-day cumulative precipitation errors were close to unbiased for lead days 0–4, but showed a systematic underestimation for lead days 5–9 in Herent (ME =−7.5 mm). Variability was substantial in both periods, with individual events exhibiting errors up to 35 mm. While several forecast periods showed good agreement, severe underestimation occurred in some cases with high rainfall totals, highlighting the limited predictability of intense precipitation events at longer lead times.

Data availability

Data will be made available on request.

References

Allen, R.t.d.G., Pereira, L.S., Raes, D., Smith, M., Ab, W., 1998. Fao,1998. Irrig. Drain. Pap. No. 56, FAO http://dx.doi.org/10.1016/j.eja.2010.12.001.

Alves, R.G., Souza, G., Maia, R.F., Tran, A.L.H., Kamienski, C., Soininen, J.P., Aquino, P.T., Lima, F., 2019. A digital twin for smart farming. In: 2019 IEEE Global Humanitarian Technology Conference. GHTC 2019, Institute of Electrical and Electronics Engineers Inc., http://dx.doi.org/10.1109/GHTC46095.2019.9033075.

Boas, T., Bogena, H.R., Ryu, D., Vereecken, H., Western, A., Hendricks Franssen, H.J., 2023. Seasonal soil moisture and crop yield prediction with fifth-generation seasonal forecasting system (SEAS5) long-range meteorological forecasts in a land surface modelling approach. Hydrol. Earth Syst. Sci. 27 (16), 3143–3167. http://dx.doi.org/10.5194/HESS-27-3143-2023.

Brinkhoff, J., Hornbuckle, J., Ballester Lurbe, C., 2019. Soil moisture forecasting for irrigation recommendation. IFAC-PapersOnLine 52 (30), 385–390. http://dx.doi.org/10.1016/J.IFACOL.2019.12.586.

Bulygina, N., Gupta, H., 2009. Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation. Water Resour. Res. 45 (2), http://dx.doi.org/10.1029/2007WR006749.

Cheng, M., Wang, H., Fan, J., Zhang, S., Liao, Z., Zhang, F., Wang, Y., 2021. A global meta-analysis of yield and water use efficiency of crops, vegetables and fruits under full, deficit and alternate partial root-zone irrigation. Agricult. Water. Manag. 248, 106771. http://dx.doi.org/10.1016/J.AGWAT.2021.106771.

El-Naggar, A.G., Hedley, C.B., Horne, D., Roudier, P., Clothier, B.E., 2020. Soil sensing technology improves application of irrigation water. Agricult. Water. Manag. 228, http://dx.doi.org/10.1016/J.AGWAT.2019.105901.

Eurostat, 2019. Agri-environmental indicator - irrigation. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agri-environmental_indicator_-_irrigation#Analysis_at_EU_and_country_level.

Goap, A., Sharma, D., Shukla, A.K., Rama Krishna, C., 2018. An IoT based smart irrigation management system using machine learning and open source technologies. Comput. Electron. Agric. 155, 41–49. http://dx.doi.org/10.1016/j.compag.2018.09.040.

Hendrickx, M.G.A., Vanderborght, J., Janssens, P., Bombeke, S., Matthyssen, E., Waverijn, A., Diels, J., 2025. Pooled error variance and covariance estimation of sparse in situ soil moisture sensor measurements in agricultural fields in flanders. SOIL 11, 435–456. http://dx.doi.org/10.5194/SOIL-11-435-2025.

Hendrickx, M.G.A., Vanderborght, J., Janssens, P., Laloy, E., Bombeke, S., Matthyssen, E., Waverijn, A., Diels, J., 2026. Field-scale soil moisture predictions in real time using in situ sensor measurements in an inverse modeling framework: SWIM2. Water Resour. Res. 62, e2025WR041324. http://dx.doi.org/10.1029/2025WR041324.

Jabro, J.D., Stevens, W.B., Iversen, W.M., Allen, B.L., Sainju, U.M., 2020. Irrigation scheduling based on wireless sensors output and soil-water characteristic curve in two soils. Sensors 20 (5), 1336. http://dx.doi.org/10.3390/S20051336.

Mehta, P., Siebert, S., Kummu, M., Deng, Q., Ali, T., Marston, L., Xie, W., Davis, K.F., 2024. Half of twenty-first century global irrigation expansion has been in water-stressed regions. Nat. Water 2 (3), 254–261. http://dx.doi.org/10.1038/s44221-024-00206-9.

METER Group, 2018. TEROS 10. URL: http://publications.metergroup.com/Manuals/20788_TEROS10_Manual_Web.pdf.

Newlands, N.K., Zamar, D.S., Kouadio, L.A., Zhang, Y., Chipanshi, A., Potgieter, A., Toure, S., Hill, H.S., 2014. An integrated, probabilistic model for improved seasonal forecasting of agricultural crop yield under environmental uncertainty. Front. Environ. Sci. 2 (JUN), 88570. http://dx.doi.org/10.3389/FENVS.2014.00017.

Raes, D., 2002. BUDGET: A soil water and salt balance model. Ref. Man. Version 5.

Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009. Aquacrop-the FAO crop model to simulate yield response to water: II. main algorithms and software description. Agron. J. 101 (3), 438–447. http://dx.doi.org/10.2134/agronj2008.0140s.

Roberts, C., Ingleby, B., Geer, G., Hólm, E., Janousek, M., Prates, F., Rodwell, M., 2024. IFS upgrade improves near-surface wind and temperature forecasts. URL: https://www.ecmwf.int/en/newsletter/181/earth-system-science/ifs-upgrade-improves-near-surface-wind-and-temperature.

Schübl, M., Stumpp, C., Brunetti, G., 2022. A Bayesian perspective on the information content of soil water measurements for the hydrological characterization of the vadose zone. J. Hydrol. 613, 128429. http://dx.doi.org/10.1016/J.JHYDROL.2022.128429.

Simunek, J., Van Genuchten, M.T., Sejna, M., 2005. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Univ. California-Riverside Res. Rep. 3, 1–240.

STATBEL, 2023. Key figures on agriculture 2023. URL: https://statbel.fgov.be/en/news/key-figures-agriculture-2023.

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 (3), 426–437.

Sui, R., 2018. Irrigation scheduling using soil moisture sensors. J. Agric. Sci. 10 (1), http://dx.doi.org/10.5539/jas.v10n1p1.

Togliatti, K., Archontoulis, S.V., Dietzel, R., Puntel, L., VanLoocke, A., 2017. How does inclusion of weather forecasting impact in-season crop model predictions? Field Crop. Res. 214, 261–272. http://dx.doi.org/10.1016/J.FCR.2017.09.008.

Torres-Sanchez, R., Navarro-Hellin, H., Guillamon-Frutos, A., San-Segundo, R., Ruiz-Abellón, M.C., Domingo-Miguel, R., 2020. A decision support system for irrigation management: Analysis and implementation of different learning techniques. Water (Switzerland) 12 (2).

Van Dam, J.C., Huygen, J., Wesseling, J.G., Feddes, R.A., Kabat, P., Van Walsum, P.E.V., Groenendijk, P., Van Diepen, C.A., 1997. Theory of SWAP Version 2.0; Simulation of Water Flow, Solute Transport and Plant Growth in the Soil-Water-Atmosphere-Plant Environment. Technical Report, DLO Winand Staring Centre.

VMM, 2025. Evolutie neerslagtekort. URL: https://vmm.vlaanderen.be/feiten-cijfers/water/droogte/indicator-neerslagtekort.

Vories, E., Sudduth, K., 2021. Determining sensor-based field capacity for irrigation scheduling. Agricult. Water. Manag. 250, 106860. http://dx.doi.org/10.1016/J.AGWAT.2021.106860.

Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. Environ. Model. Softw. 75, 273–316. http://dx.doi.org/10.1016/j.envsoft.2015.08.013.

Wang, X., Wang, X.S., Li, N., Wan, L., 2021. Bayesian inversion of soil hydraulic properties from simplified evaporation experiments: Use of DREAM(ZS) algorithm. Water 13 (19), 2614. http://dx.doi.org/10.3390/W13192614.

Wang, W., Wei, Y., Hao, L., Wei, Z., Zhao, T., 2025. Soil moisture forecasting in wireless sensor networks via spatiotemporal graph convolutional networks. Vadose Zone J. 24 (1), e70000. http://dx.doi.org/10.1002/VZJ2.70000.

Waverijn, A., Vandewoestijne, E., Janssens, P., Vaerten, J., Hisette, N., Matthyssen, E., Hendrickx, M., Vanderborght, J., Diels, J., 2024. Optimizing drip irrigation strategy in leek (allium ampeloprasum var. porrum). Acta Hortic. (1398), 185–192. http://dx.doi.org/10.17660/ACTAHORTIC.2024.1398.25.

Yang, W.H., Clifford, D., Minasny, B., 2015. Mapping soil water retention curves via spatial Bayesian hierarchical models. J. Hydrol. 524, 768–779. http://dx.doi.org/10.1016/J.JHYDROL.2015.03.029.

Yang, J., Reichert, P., Abbaspour, K.C., Yang, H., 2007. Hydrological modelling of the chaohe basin in China: Statistical model formulation and Bayesian inference. J. Hydrol. 340 (3–4), 167–182. http://dx.doi.org/10.1016/J.JHYDROL.2007.04.006.

Zinkernagel, J., Maestre-Valero, J.F., Seresti, S.Y., Intrigliolo, D.S., 2020. New technologies and practical approaches to improve irrigation management of open field vegetable crops. Agricult. Water. Manag. 242, 106404. http://dx.doi.org/10.1016/J.AGWAT.2020.106404.