Linking soil moisture sensors and crop models for irrigation management

Antoine Haddon, Loïc Kechichian, Jérôme Harmand, Cyril Dejean, Nassim Ait-Mouheb

Linking soil moisture sensors and crop models for irrigation management

Antoine Haddon*[a], Loïc Kechichian[b], Jérôme Harmand[a], Cyril Dejean[b], Nassim Ait-Mouheb[b]

[a]INRAE, Univ Montpellier, LBE, 102 Avenue des Etangs, Narbonne, France.
[b]INRAE, Univ Montpellier, UMR GEAU, Montpellier, France.
*Corresponding author (antoine.haddon@inrae.fr)

Abstract
A number of challenges must be faced when using soil moisture sensors, such as accounting for soil heterogeneity in measurements or dealing with sensor faults. As a consequence, it is difficult to obtain reliable estimations of the water status in the root zone and using sensor data for irrigation planning is not straightforward. In this work, a method is proposed to interpret soil water content measurements that is based on the use of a model to correct and complement sensor data, in particular in the case of a non uniform water distribution. This approach relies on the assumption that porosity is the main driver of heterogeneity in hydraulic properties at small scales, which allows to factor out the spatial variations of the sensor's signal. With practical applications in mind, a simple model and an efficient calibration procedure are developed, in particular considering the online application of the method to complement sensor data in real time. The capabilities of the model are illustrated with data from experiments on the growth of lettuces in greenhouses with reclaimed wastewater irrigation. Requiring only a short calibration period, the model is successfully validated and is proven to be a valuable tool to correct for sensor malfunctions. Moreover, the proposed method is shown to allow the meaningful estimation of the water status of the soil crop system, in particular when measurements of sensors positioned close to each other showed important differences.

Keywords: soil water sensor ; simplified model ; irrigation scheduling

1 Introduction

Increasing water scarcity and the essential role irrigation plays in food security in many regions of the world, calls for a better management of water uses in agriculture. A key element of any efficient irrigation system that is capable of achieving high yields, is the appropriate planning of water inputs to meet the crops needs. This requires answering two questions : when irrigation should be triggered and how much water needs to be applied. To help in taking these decisions, different approaches have been used and often rely on information on the water status and dynamics of the soil-crop system either through the use of sensors or with models incorporating weather data (Abioye 2020, Villalobos 2016).

Sensors have been developed that allow the measurement of the soil water content (SWC) or the water potential and have been used independently of models for irrigation planning. Typically, in this case, the decision to irrigate is taken when a certain threshold is reached, which represents the onset of plant water stress. Then, to avoid excess irrigation and compute the appropriate irrigation volume it is necessary to also know the soil field capacity. However, determining the irrigation threshold value and the field capacity for a given soil remains challenging (Vories 2021). Appropriate sensor calibration also represents a difficulty, with various effects impacting sensors, such as temperature, that must be compensated (Feng 2020). Cheaper sensors have been introduced to reduce costs but these are less reliable, with sensor faults that lead to corrupted data, further complicating the use of sensors for irrigation (Bogena 2007). Nonetheless, among the issues that have been reported with using sensors for irrigation scheduling, the interpretation of sensor data has been considered as the main difficulty (Sui 2020). In particular, a number of problems arise from a non uniform distribution of water in the soil, since spatial variations in SWC makes it challenging to determine if the sensor output actually represents the water status of the root zone. As a result, the positioning of sensors has an important impact on measurements, with differences observed in sensors over short distances (Bogena 2007, Vories 2021).

Approaches for irrigation planning based on models include water balance methods, which use weather data along with a model to estimate evapotranspiration losses, and compute irrigation to compensate (Allen 1998, Villalobos 2016, Pereira 2020). From the first simple models, more complex dynamic models have been developed, with detailed soil water balances to estimate the distribution of water in the soil column (Brisson 2003, Rodríguez-Iturbe 2007, Mailhol 2011, Cheviron 2016, Šimůnek 2018). However, a more detailed representation of the water dynamics leads to increased complexity and models that are harder to use, with calibration of model parameters representing a serious difficulty for the most complex models. This is an important consideration for a model that must be used in practice, as a model must first be calibrated to be applied to each specific situation.

Combining both models along with sensors is an interesting solution to overcome the shortcomings of each approach when used separately. On one hand, sensors offer the possibility to link models to reality, with sensor data used for the calibration of models. On the other hand, models can be used to correct sensor output in the case of faults but also allows the integration of weather and sensor data to gain a broader perspective and a better understanding of a specific irrigation problem. In addition, this opens the possibility for the use of a wide variety of tools from monitoring and control engineering for irrigation management (Abioye 2020, Cobbenhagen 2021). However, this raises the often overlooked issue of linking sensors and models and how to assimilate data from sensors into models. In particular, due the difficulties in interpreting sensor data, it is not straightforward to establish a correspondence between sensor output and model variables. In this work, we focus on the problems associated with the non uniform distribution of soil water content, and how to deal with soil heterogeneity when using sensor data along with a model. In section 2, we present a model, explaining the link between sensor output and model variables and then we detail a parameter calibration procedure specifically developed. This approach is tested with experimental data, presented in section 3. Results are discussed in section 4 before conclusions in section 5.

2 Model Presentation

2.1 Soil water content

Variations in the distribution of water in the soil are caused by a number of factors, which have impacts at different scales. In an area exposed to similar conditions, differences in SWC can be the result of heterogeneities in soil composition and structure (Warrick 2001).

Soil composition, i.e. mineral particle size distribution and organic matter content, can vary at the scale of a field but locally, such as near a sensor, it is reasonable to consider it constant. Spatial variations in SWC can be the result of heterogeneities in the soil structure, in particular the porosity, and this can occur at a small scale. Indeed, differences in soil compaction will lead to variations in the distribution of pores, affecting the water holding capacity and the entire soil water dynamics.

In this study, the measurement of SWC from sensors based on capacitance and frequency domain reflectometry technology are considered. These rely on the relation between the SWC and the dielectric permittivity of the soil due to the different permittivities of water, minerals and air. In turn, capacitance is related to the permittivity of the medium surrounding the sensors, and therefore capacitance can be used to indirectly measure SWC. The output of these sensors is then converted through a regression curve to the volumetric soil water content (VWC), which is the ratio of the water volume relative to the total volume :

$$\theta = \frac{WaterVolume}{TotalVolume} \tag{1}$$

The total volume of soil can be decomposed into the volume of solid components and the pore volume, which itself is composed of the volume of air and water. As a consequence, porosity will thus impact the sensor's measurements and in particular, a spatial heterogeneity of porosity will result in variations of VWC. Typically, if two sensors are positioned close by in an area where the soil water status should be the same, the sensors can output different measurements due to small variations of porosity. This raises the issue of interpreting sensor data and in particular the validity of a measurement for a given area.

As a consequence, using sensor data along with a model for irrigation planning requires some methodology to account for the local spatial variations in soil water measurements. One possibility is to use a model with variables for soil water expressed as VWC and to represent soil heterogeneity in detail. This is the approach followed by many dynamic soil and crop models, and this leads to considering a fine representation of the soil to compute in detail the distribution of water (Brisson 2003, Šimůnek 2018). However, with this approach, soil heterogeneity is represented by varying the parameters associated with the different soil properties. Indeed, the characteristic moisture levels, such as field capacity or wilting point, when expressed in VWC, would have to be different at various points in space to reflect the variations of soil structure. In particular, this raises the issue of the calibration of these distributed parameters, which are functions of space. At least, it would be necessary to calibrate, for each sensor position, all the parameters representing characteristic soil moisture levels, such as field capacity and wilting point.

The complexity of this approach and the associated models can be questioned in practical applications, where model usability and efficiency are important. Furthermore, models for decision support do not necessarily need to be as detailed as those developed for scientific purposes, in particular if they are to be used by farmers. For the problem studied here, the limited online measurements from a few sensors and the little information on the soil properties available in practice make it difficult to accurately calibrate complex models and justify the use of simpler models.

Instead, the model and the associated calibration method proposed here consider soil water variables expressed as Pore Water Content (PWC), which is the soil water volume relative to pore volume :

$$S = \frac{WaterVolume}{PoreVolume} \quad (2)$$

Porosity ϕ is defined as the ratio of pore volume to total volume,

$$\phi = \frac{PoreVolume}{TotalVolume} \quad (3)$$

Therefore, PWC and VWC are related by

$$\theta = \phi S \quad (4)$$

Then, sensor data, which is recorded as VWC, can be readily converted to PWC once porosity has been estimated.

The choice of using PWC variables is based on the hypothesis that the local variations in soil water content are the result of local differences in soil porosity. Furthermore, we make the assumption that the VWC measurements can be decomposed according to (4) into the product of porosity, which represents the soil's spatial heterogeneity, and PWC, which is assumed to be constant locally. In other words, we suppose that the sensor output $\theta(t, x)$ which varies in space and time, is the product of a term which varies only in space, porosity $\phi(x)$, and another term which varies only in time, PWC $S(t)$:

$$\theta(t, x) = \phi(x)S(t) \quad (5)$$

with x representing a space coordinate and t the time coordinate.

An important advantage of this approach is that it does not require a fine representation of the soil and instead the variables of PWC will represent the water content in a large area where soil type is constant. This results in an efficient model with only a few variables. Furthermore, the porosity is the only parameter used to account for differences in recorded VWC and is calibrated to the sensor data. Indeed, it is assumed that, when expressed as PWC, the characteristic soil moisture levels depend only on the soil type but not on the porosity. Thus, considering that the soil type is homogenous in an area results in constant characteristic soil moisture levels. The associated parameters can therefore be set from reference values based on the soil composition. This method has the advantage of requiring the calibration of only one parameter per sensor to explain the spatial variations and this is important in obtaining an efficient and generic calibration method.

To justify these hypotheses, first note that characteristic soil moisture levels, when they are expressed as levels of soil water potential, can be considered to be independent of soil porosity or composition (Laio, 2001). For example, the wilting point is assumed to correspond to a soil water potential of -3 MPa, with variations essentially due to plant type but not due to soil composition or porosity. Next, a soil water retention curve can be used to convert water potential to VWC and in particular the following relation has been proposed (Clapp 1978),

$$\left(\frac{\psi}{\psi_s}\right)^b = \frac{\theta}{\phi} \quad (6)$$

Here, ψ is the soil water potential, ψ_s and b are parameters depending on the soil type. The water retention curve therefore encapsulates the different sources of spatial variation in water content due to soil type and porosity. Note that VWC and PWC are related through porosity and can be rewritten to relate the soil water potential to PWC:

$$\left(\frac{\psi}{\psi_s}\right)^b = S \quad (7)$$

The important observation here is that this relation is independent of porosity. The consequence is that, if the characteristic soil moisture levels correspond to constant soil water potentials, then they do not vary with porosity when they are expressed as PWC. For example, assuming that the soil water potential at the wilting point and the soil type are constant in a given area implies that the wilting point is also constant when expressed in PWC. Note however that the parameters ψ_s and b vary with soil composition and thus this gives the dependence on soil type of the PWC characteristic soil moisture levels.

2.2 Model dynamics

The model proposed here is a soil water balance compartment model, each compartment corresponding to a layer of soil in an area around a sensor. The main variables are the soil water content, expressed in PWC, and are assumed to represent soil water content of the area surrounding each sensor where the soil composition is constant. The horizontal movement of water is neglected and therefore areas of different soil composition are not connected. However, to represent vertical variation and movement of water, several layers can be considered as a series of interconnected compartments. The number of layers depends on the number of sensors used and can be adapted to the crop and soil type, with the case of crops with deep roots or vertically heterogeneous soils requiring more layers. The model developed here considers a division of the soil column in 2 layers, with the objective of using data from sensors positioned at 2 different depths.

The dynamics are obtained by considering the balance of inputs and outputs in each compartment. Unlike many crop models that use a fixed time step, the model presented here considers a continuous time, which is better suited to capture phenomena with different timescales such as those present in the soil water dynamics. The result is a continuous dynamical system described by a set of ordinary differential equations. The choice of this approach also allows to take advantage of the tools developed in dynamical systems theory, such as observer and control methods (Khalil, 2015).

To compute losses due to evapotranspiration, crop growth is computed and the model presented here uses concepts originated in the methods of the FAO Irrigation and Drainage Paper No. 56 (Allen 1998) and further developed in the AquaCrop model (Steduto 2009). Such concepts have already been applied to construct continuous dynamical systems crop models (Laio, 2001, Rodríguez-Iturbe 2007, Pelak 2017). The model presented here largely follows these works but with a few modifications. In particular, only vertically homogeneous soils had been considered and here these models are extended to the multi-layer case.

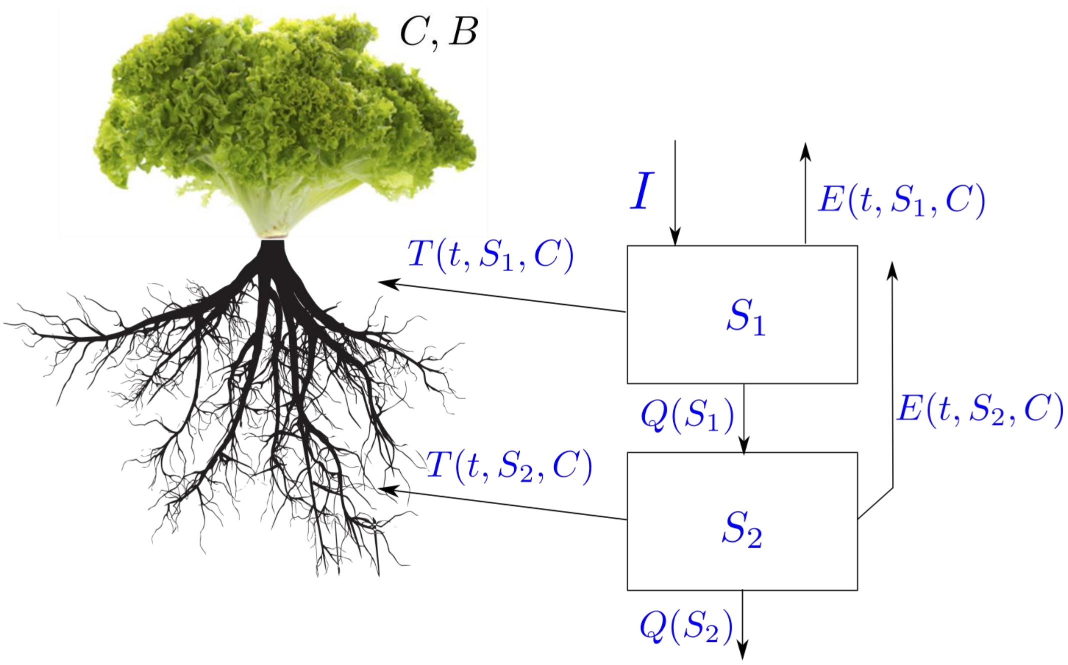

Figure 1: Schematic representation of the 2 layer model. $S_i$is the pore water content (PWC) of layer i, C canopy cover, B biomass, I irrigation, E evaporation, T transpiration and Q leakage.

The model variables are the PWC in the top layer S_1, and in the bottom layer S_2. The crop is represented with the above ground dry biomass B [kg m^{-2}] and canopy cover C, which is the fraction of ground shaded by the canopy.

Denoting Z_i the height and ϕ_i the porosity of the layer i (with i = 1 or 2), then $\phi_i Z_i$ represents the active depth, *i.e.* the volume per unit area of pore space in the considered layer. Therefore $\phi_i Z_i S_i$ is the height of water of the layer and is the quantity on which the balance is written, considering input from irrigation I and losses due to transpiration T, evaporation E and drainage Q.

$$\phi_1 Z_1 \frac{dS_1}{dt}(t) = I(t) - T(t, S_1, C) - E(t, S_1, C) - Q(S_1) \quad (8)$$

$$\phi_2 Z_2 \frac{dS_2}{dt}(t) = Q(S_1) - T(t, S_2, C) - E(t, S_2, C) - Q(S_2) \quad (9)$$

The soil water balance follows the dual crop coefficient method and uses the canopy cover to partition the Penman-Monteith reference evapotranspiration ET_0 into transpiration and evaporation. Then the transpiration flux, from layer i = 1 or 2, is computed as:

$$T(t, S_i, C) = K_{cb}^i K_S(S_i) C ET_0(t) \quad (10)$$

with$K_{cb}^i$the crop transpiration coefficient and K_S the water stress function.

$$K_S(S) = \begin{cases} 0 \text{ for } S \leq S_W \\ \frac{S-S_W}{S_*-S_W} \text{ for } S_W < S \leq S_* \\ 1 \text{ for } S_* < S \end{cases} \quad (11)$$

with S_W the wilting point and S_* the water stress level.

For simplicity and to obtain a parsimonious model, several processes are not explicitly represented in this model, as for instance root growth. Limited root growth could limit transpiration, in particular in the early stages of plant life, but it is considered here that this effect is indirectly taken into account through the presence of the canopy cover in the expression of transpiration which already limits the crops' water consumption. In addition, a difference in root density between the 2 layers could change the transpiration flux from each layer but this is accounted for by taking different crop transpiration coefficients K_{cb}^i in each layer.

Similarly, the evaporation flux from layer i = 1 or 2 is computed as:

$$E(t,S,C) = K_e^i K_r(S)(1-C)ET_0(t) \quad (12)$$

with K_e^i the evaporation coefficient and K_r the evaporation reduction function.

$$K_r(S) = \begin{cases} 0 \text{ for } S \leq S_h \\ \frac{S-S_h}{1-S_h} \text{ for } S_h < S \leq 1 \end{cases} \quad (13)$$

The two layers are connected through the leakage term, with the water draining from the top layer $Q(S_1)$ feeding into the bottom layer. Leakage is modelled with a tipping bucket approach, with no flow when PWC is less than field capacity S_{fc},

$$Q(S) = \begin{cases} 0 \text{ for } S \leq S_{fc} \\ k_{sat}\frac{S-S_{fc}}{1-S_{fc}} \text{ for } S_{fc} < S \leq 1 \end{cases} \quad (14)$$

with k_{sat} the saturation conductivity.

A logistic equation is considered for the canopy cover and the growth rate is proportional to crop transpiration, to account for limitations in case of water stress.

$$\frac{dC}{dt}(t) = r_G\big(T(t,S_1,C) + T(t,S_2,C)\big)\left(1 - \frac{C}{C_{max}}\right)$$
(15)

with r_G the potential canopy growth rate and C_{max} the maximum canopy cover.

Biomass growth follows the concept of water productivity used in Aquacrop, with growth proportional to the total transpiration flux,

$$\frac{dB}{dt}(t) = W_* \frac{T(t,S_1,C)+T(t,S_2,C)}{ET_0(t)} \tag{16}$$

with W_* the daily water productivity.

The model considered is thus composed of equations (8), (9), (15), (16) and schematically represented in Figure 1.

The model can be used to compute the total water losses due to evaporation, transpiration and leakage over a given time interval [0,T]:

Total Evaporation : $\int_0^T E(t,S_1,C) + E(t,S_2,C)dt$

Total Transpiration : $\int_0^T T(t,S_1,C) + T(t,S_2,C)dt$

Total Leakage : $\int_0^T Q(S_2)dt$

2.3 Parameter calibration

The model parameters can either be set from reference values from public databases, directly measured or estimated from available sensor data. The latter case requires solving the optimisation problem of minimising the error between model simulations and measurement data. This can be computationally intensive and present a number of challenges, such as the problem of identifiability (Walter 1997). For these reasons, it is important to set as many parameters as possible by other means to obtain an efficient calibration.

Table 1 lists the parameters set from references. As previously explained, the parameters representing soil hydraulic properties (characteristic soil moisture levels S_h, S_w, S_*, S_{fc} and the saturated conductivity k_{sat}) values are selected based on the soil type which can be identified with granulometric measurements. Values of these characteristic soil moisture levels in PWC for different soil types can be found in (Laio, 2001). The parameter C_{max} represents the maximum area of the soil surface that can be shaded by the crops canopy and therefore depends on the plant type as well as geometric consideration, such as row spacing.

The height Z_i of each compartment can be set by first considering that the modelled soil layers correspond to a compartment where the water content or soil type does not vary considerably. In addition, sensors have a given volume of influence and measurements correspond to an average over this volume. Therefore, with the objective of reproducing sensors data, the heights Z_i should be related to the vertical size of the volume of influence of the soil moisture sensors. Accordingly, knowledge of the soil column composition, its variation and the properties of the sensors should be taken into account in positioning sensors and setting the height parameters Z_i.

All other parameters (Table 2) are calibrated by minimising the error between simulations and measurement data. The root mean square (RMSE) is used to compute errors :

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(D_k - X(t_k)\right)^2}$$

(17)

with N the number of data points, D_k the measured data at time t_k and $X(t_k)$ the simulated value. When combining several variables expressed in different units (i.e. PWC and canopy cover), errors are computed by summing the relative RMSE (relRMSE) for each variable :

$$relRMSE = \frac{RMSE}{\bar{D}} \tag{18}$$

with $\bar{D}$ the mean of the data.

The porosity ϕ_i is the main parameter that must be calibrated as it is considered here to account for the spatial variability of soil properties. Several sets of sensors can be used to monitor the soil for redundancy purposes or to study variations between different areas. In these cases, the porosity must be estimated for each soil water sensor. When sensors are used over several growth cycles, if the soil surrounding the sensors is left undisturbed, the same parameters can be used for a new growth cycle. However, if the sensor is positioned in proximity to roots or if the soil is disturbed due to tillage between growth cycles, then the soil structure will change and thus porosity must be estimated again.

The other calibrated parameters are independent of individual sensor's and it is possible to use the same parameters for different sets of sensors. However, the use of the same set of parameters over several production cycles or if the growing conditions are different is limited. Indeed, due to the simplicity of the model considered here, a number of effects are not taken into account and instead the impact of processes not represented end up hidden in parameters. For example, the impact of temperature on crop growth is only taken indirectly into account through reference evapotranspiration but it has been known for a long time that biomass growth depends on temperature, with the concept of growing degree days. For this model, this means that the biomass growth rate in fact depends on temperature. As a consequence it is necessary to re-calibrate parameters for different production cycles or if growing conditions change, depending on the role of each parameter. Table 2 lists the circumstances for which each parameter must be calibrated.

In practice, due to the importance of the porosity on the water dynamics, this parameter can be first estimated alone, to get a preliminary ajustement of the general features of the soil water dynamics. Then, the precise calibration is conducted in a second step, with the estimation at the same time of the porosity along with the evapotranspiration and canopy growth rate parameters to get a precise fit for the S and C variables. This is the step that is the most challenging, as it can require the estimation of up to 7 parameters and thus there is a strong interest in reusing parameters from a previous calibration if possible. Finally, as the biomass does not affect the dynamics of other variables in the model presented here, the growth rate (W_*) can be estimated independently at the end to obtain a good adjustment of the biomass.

Table 1 : Model parameters, set from references

Parameter	Value	Units	Name	Source
Z_i	100	mm	Depth of layer i	METER Group
S_h	0.19	-	Hygroscopic point	Laio, 2001
S_W	0.24	-	Wilting point	Laio, 2001
S_*	0.57	-	Point of incipient stomatal closure	Laio, 2001
S_{fc}	0.65	-	Field capacity	Laio, 2001
k_{sat}	200	mm/d	Saturated hydraulic conductivity	Laio, 2001
C_{max}	0.8	-	Maximum canopy cover	-

Table 2 : Model parameters, calibrated from data.

Parameter	Units	Name	Calibration
ϕ_i	-	Porosity	For each sensor
K_{cb}^i	-	Transpiration crop coefficient	For each production cycle
K_e^i	-	Evaporation crop coefficient	For each soil type
r_G	1/d	Canopy cover growth rate	For each production cycle and growing condition
W_*	kg m²/d	Normalised daily water productivity	For each production cycle

3 Calibration and validation data

The approach presented here is illustrated in the context of experiments in wastewater reuse, in which irrigation with freshwater and reclaimed water is compared. This offers the possibility to showcase the use of a model for the interpretation and correction of sensor measurements. Furthermore, this allows to demonstrate the capabilities of the model and the associated calibration procedure in different growing conditions. It should be noted that the quantitative control of treated wastewater reuse in agriculture is an important issue, considering that it is a resource that can be limited and moreover to avoid possible sanitary and agronomic impacts of uncontrolled wastewater irrigation (Ait-Mouheb et al. 2018).

3.1 Experimental site

The experimental site is located at Murviel-lès-Montpellier, in the south of France (43.605° N 3.757° E), on a wastewater treatment plant which is equipped with a constructed wetland, composed of reed bed filters with forced aeration, and with additional secondary treatment with ferric chloride as flocculant to remove phosphorus. Two greenhouses of 100 m² each have been in use since 2017, to run experiments on the impact of wastewater reuse in agriculture (Ait-Mouheb et al. 2022). Large soil bins (1m² and 60 cm soil depth) are used to isolate the experiments and avoid field contamination resulting from irrigation with reclaimed wastewater. The bins were filled with loamy clay soil (24.5% clay, 32% fine silt, 13.7% silt, 10.6% fine sand and 19.2 % of sand).

Lettuces (*Lactuca sativa)* were grown in 2021, with 8 plants per bin and starting with plantlets at the 3 leaf stage. Two growth cycles of 6 weeks were conducted, from 13 April to 25 May and from 27 May to 5 July, hereafter referred to as cycle 1 and cycle 2, respectively. Meteorological variables were measured with a weather station located in the greenhouse. Hourly air temperature, relative humidity and global radiation levels were recorded during both growth cycles.

3.2 Irrigation and fertilisation

The bins were irrigated with different water qualities and in this study we focus on the 2 bins in which soil moisture sensors were positioned, with one bin irrigated with freshwater (FW) and another with treated wastewater (TWW). Drip irrigation was conducted with one surface dripper per lettuce and one dripper without any plant in the centre of the bin. The drippers deliver a nominal flow rate of 2 L/h and flow rates were monitored during the irrigation cycles and showed no significant variation. According to the manufacturer's recommendations, irrigation water was filtered at 130 μm before irrigation to prevent physical clogging of the drippers.

Irrigation was performed twice a week during cycle 1 and 3 times per week during cycle 2, with irrigation volumes computed to compensate for evapotranspiration and guarantee a VWC above 0.15. Evapotranspiration was estimated using the method from the Food and Agricultural Organization of the United Nation (FAO) Irrigation and Drainage Paper No. 56 (Allen 1998). First, reference evapotranspiration ET_0 was computed from weather data with the Penman-Monteith equation. Then, evapotranspiration was computed as:

$$ET = K_c ET_0 \tag{19}$$

using a crop coefficient $K_C = 0.4$ from germination to 18 leafs and then $K_C = 0.8$ from 18 leaves to harvest (Berry, 2013).

The soil of each bin was analysed to determine the available nitrogen (N), phosphorus (P) and potassium (K) for crop growth at the beginning of each growth cycle. In addition, as these nutrients are present in wastewater, the amount of N, P and K supplied through TWW irrigation was estimated from analysis of the treated wastewater and typical irrigation requirements. Then considering the needs of the lettuce (Berry, 2013), the nutrients already present in the soil and the possible contribution from irrigation in the TWW bin, fertilisation was conducted to provide for the needs of the lettuce over a cycle and ensure the same level of nutrients in the TWW and FW bins. Accordingly, N was supplied to the FW bin for cycle 1, P for the TWW and FW bin for cycle 2 and K for the TWW bin cycle 2.

3.3 Soil Moisture Sensors and Lettuce Growth Monitoring

For the monitoring of soil water content (SWC) 16 capacitive soil moisture sensors were installed (12 sensors of model ECH20 EC5 4 sensors of model ECH20 10HS, all from METER Group). Sensors were installed in pairs, at two different depths of 5cm and 20cm. The sensors were positioned either under a dripper with lettuce or under a dripper without a lettuce (Figure 2).

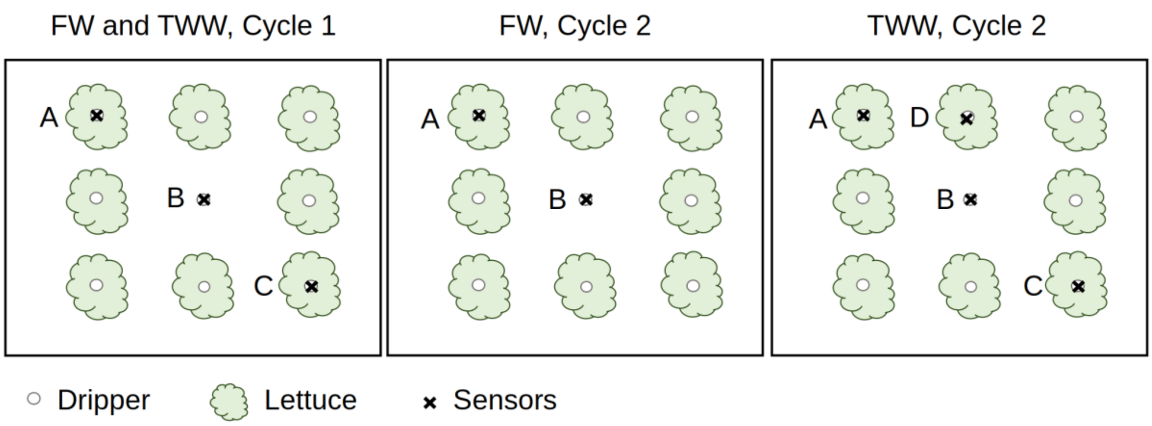

Figure 2 : Position of Drippers, Lettuces and sensors in bins irrigated with freshwater (FW) and treated wastewater (TWW).

The sensors were calibrated for the soil of the experiment, according to manufacturer instructions, to obtain the parameters to convert the output of the sensors to volumetric soil water content. The sensor data was furthermore corrected to account for the effect of temperature according to the method proposed by (Cobos, 2007).

Once a week, photos were taken of the bins and using the ImageJ software, the horizontal surface area of each lettuce was measured. Fresh and dry biomass were measured at the end of each growth cycle.

3.4 Model calibration and validation

The data from each growth cycle was separated into calibration and validation sets, with parameters estimated with data from the 3 first weeks (days 0 to 21 for cycle 1 and days 0 18 for cycle 2) and validated with data from the 3 last weeks (days 22 to 42 for cycle 1 and days 19 to 39 for cycle 2). For calibration and comparison between model and sensor output, the data from the soil moisture sensors was converted from VWC to PWC with the models' porosity parameter and then was compared to S_1 for sensors at 5cm depth and to S_2 for sensors at 20 cm depth. The weekly lettuce surface area was converted to canopy cover data by dividing the surface area of each lettuce by the soil surface area, considering that each lettuce was on a square of 1/9th of the soil bins area (i.e. 1111 cm^2). The canopy cover variable C of the model was compared with the average over each soil bin of the canopy cover data. In addition, the average over each bin of the final dry weight of the lettuces was compared to the final value of the model variable B.

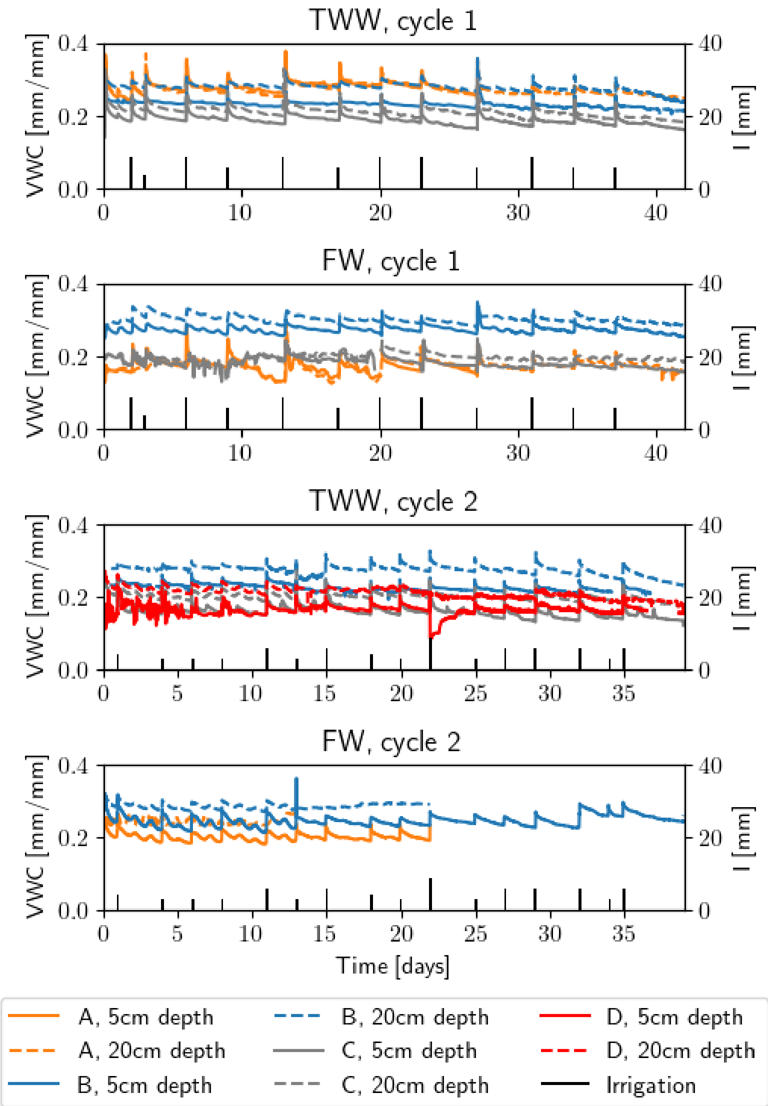

Figure 3 : Data from soil moisture sensors for cycle 1 and 2, volumetric soil water content (VWC) [mm/mm] and irrigation volume (I [mm] from treated wastewater (TWW) and freshwater (FW) irrigation. A,B,C,D denote sensor positions with A,C and D under a dripper with lettuce and B under a dripper without lettuce.

4 Results and Discussion

4.1 Soil Water Content Measurements

The volumetric soil water content as measured by the soil moisture sensors is presented in figure 3, for growth cycles 1 and 2 and for irrigation with TWW and FW. A number of sensor faults can be observed, such as problems with the recording of unphysical oscillations (sensors C, FW, cycle 1; sensors D TWW, cycle 2). Furthermore, a number of sensors malfunctioned and either recorded intermittently or produced no recording at all (sensors A, TWW, cycle 1; sensors A and B, FW cycle 2). Nonetheless, this is expected with the type of inexpensive sensors used here and has been reported as a major issue when working with cheaper sensors (Bogena 2007).

The sensor's signals follow a general pattern and two dynamics can be identified : first a fast response during irrigation events and secondly, between irrigations, a slower dynamic characterised by a steady decrease in VWC. During an irrigation, the sensor's signals exhibit spikes corresponding to a temporary water saturation of the soil followed by a rapid decrease due to drainage towards lower layers of the soil. Then, a few hours after irrigation, the SWC stabilises at a value expected to represent field capacity. The second phase takes place over the course of the several days between irrigations, with evapotranspiration driving water losses.

The amplitude of the spikes recorded during irrigations varies, even when the same irrigation volume is applied. For example, the sensor at position A, 5cm depth, in the TWW bin during cycle 1, on day 6 jumps from 0.268 to 0.35 and stabilises at 0.283 but on day 13 jumps from 0.263 to 0.378 and stabilises at 0.301, despite both days receiving 9mm of irrigation. The variations of the maximum reached can be explained by differences in the sampling time during the irrigation event. Indeed, the sampling period was of 1 hour whilst irrigations lasted 30 minutes and therefore the measurement of the maximum reached depends on whether sensors recorded during or slightly after irrigation. However, the differences in the values at which VWC stabilises afterwards, i.e. field capacity, can not be as easily explained (Vories 2021).

In between irrigations, during the phase of slower VWC decrease, small oscillations can be observed, most notably on the signal from sensors at position B in the FW bin or sensors at position C in the TWW bin. Although the amplitude varies among sensors, the period is always of one day with a minimum at early morning and a maximum at the end of the day. This indicates that this is due to daily variations in physical conditions, such as temperature which has been reported to impact measurements (Bogena 2007, Cobos 2007).

When comparing signals of the sensors within the same bin, differences can be observed in the VWC measured, despite sensors being close and exposed to the same conditions. In addition, for certain signals, these differences appear almost constant throughout the 6 weeks of each growth cycle, such as for sensors at position A and C in the TWW bin during cycle 1. Furthermore, differences can be observed in the values recorded for sensors at the same location but at different depths, for almost all positions, with SWC always higher in the bottom soil layer and again, certain sensor signals seem to differ only by a constant value (sensor C, TWW, cycle 1 and 2).

More water can be retained in the deeper layers of soil as it is less affected by evapotranspiration than the topmost layer and thus a higher SWC at 20 cm depth can be expected some time after an input of water. However, shortly after an irrigation event, the moisture content should rapidly stabilise at field capacity and therefore, if the soil had an homogeneous water holding capacity, sensors at different locations and depths should record the same SWC after irrigation.

An explanation for the differences observed could be a hardware problem, such as a sensor fault, or the result of actual spatial variations of the soil moisture distribution. It is not possible to determine from the available measurements which of these two is the cause of the observed differences and in fact it is likely a combination of both. Soil heterogeneity can lead to spatial variations in VWC (Warrick 2001) but here the soil composition can be supposed constant, especially considering the controlled conditions of the experiment, where efforts were made to render the soil as homogeneous as possible before each cycle. The local variations observed could however be explained by heterogeneities in the soil structure, as previously explained, with small differences in soil compaction leading to variations in porosity and thus affecting the soil water dynamics. In particular, the constant differences observed between certain sensor signals indicates that a factor independent of time is responsible for the spatial variations in VWC. This has led us to the suggested decomposition of the measured VWC into the product of two factors, with porosity accounting for spatial variations and PWC representing the time evolution of soil water status.

The differences between recordings within the same bin and the various sensor faults make it difficult to use such measurements directly for irrigation planning. Indeed, it is not straightforward how the sensor data can be used to estimate the water holding capacity or the soil water status at a given moment, which are both important to decide how much and when to irrigate. This justifies the use of a model as suggested here, to correct the problems with the sensor data and with the objective of irrigation decision support.

4.2 Model calibration and validation for soil water status estimation

A first calibration, using only the first 3 weeks of data of each cycle, was designed to test the performances of the model, in a context of irrigation planning. Calibration and validation errors are presented in figures 4 and comparisons of model simulations and data are shown in figures 5 and 6.

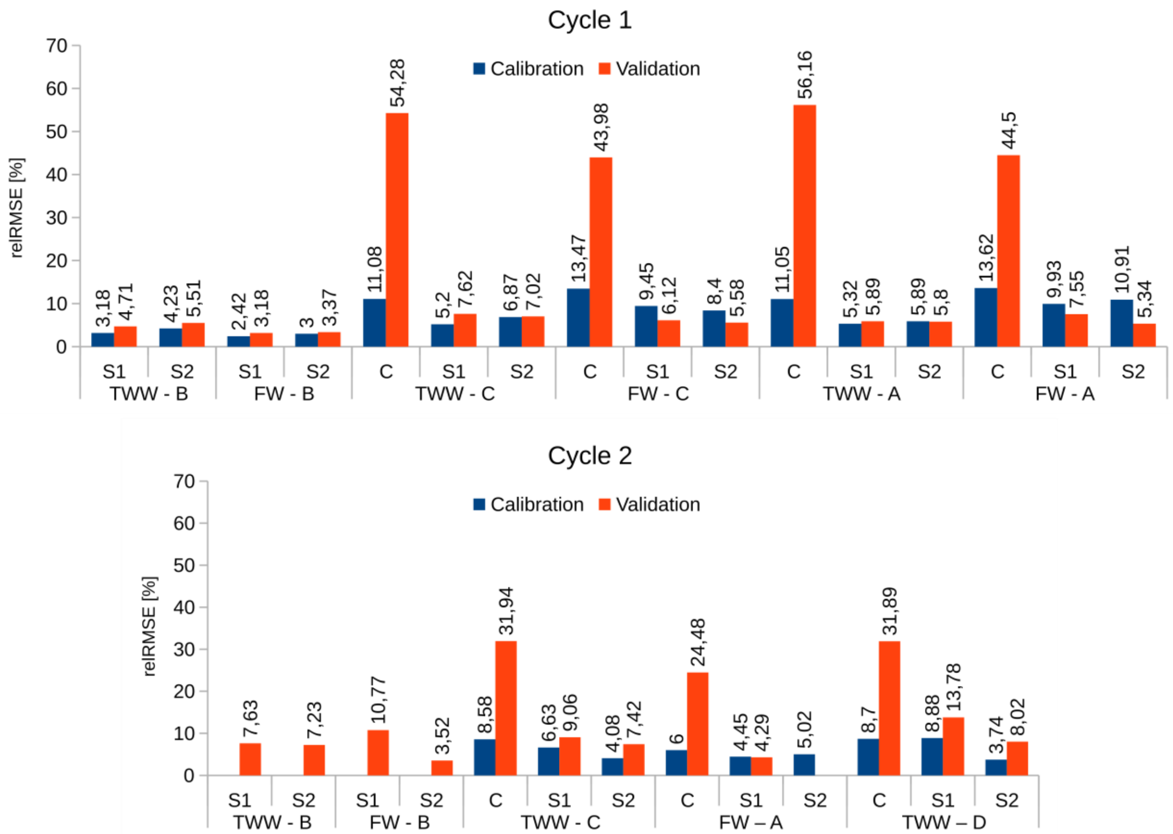

Figure 4 : Calibration and validation relative root mean square errors (relRMSE) for cycle 1 and 2, treated wastewater (TWW) and freshwater (FW) irrigation. A,B,C,D denote sensor positions with A,C,D under a dripper with lettuce and B under a dripper without lettuce.

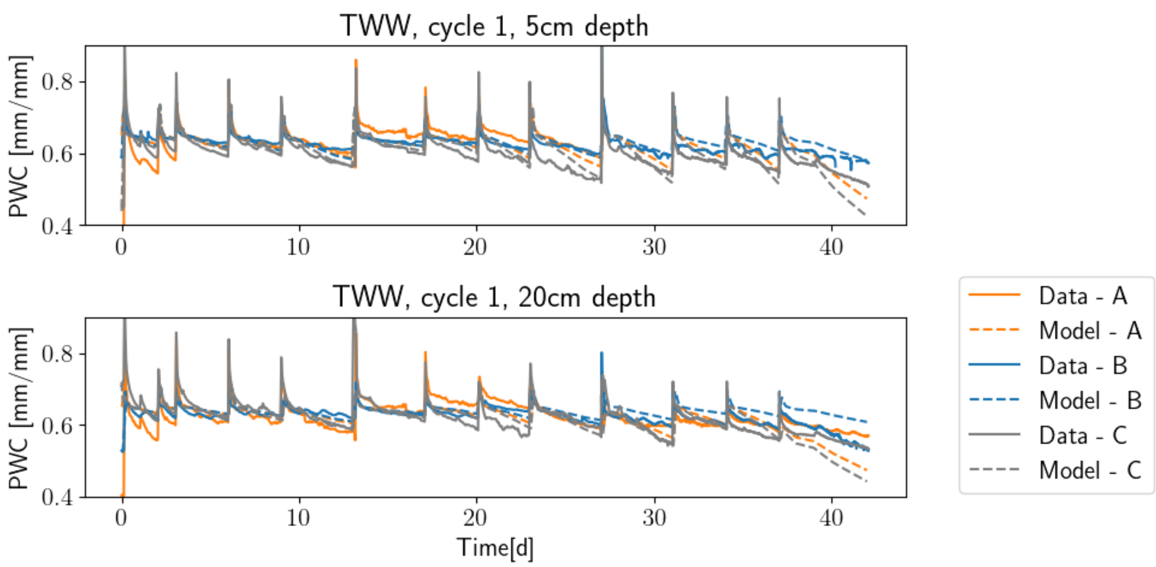

Figure 5 : Comparison of sensor data and model for the treated wastewater (TWW) bin, cycle 1. Data is converted to pore water content (PWC) with the model's porosity parameter. A,B,C denote sensor positions with A,C under a dripper with lettuce and B under a dripper without lettuce.

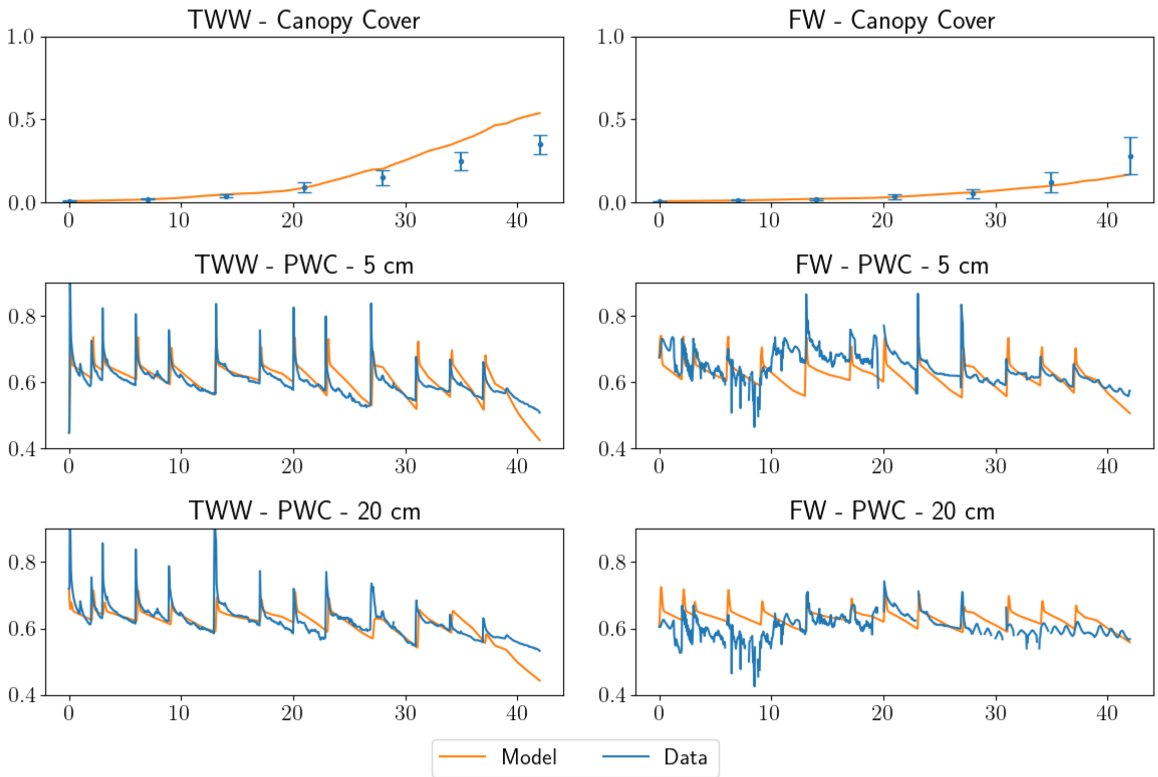

Figure 6: Comparison of data and model. Left : treated wastewater, cycle 1, sensor position C. Right : freshwater, cycle 1, sensor position A. Data for soil water is converted to pore water content (PWC) with the model's porosity parameter. Model parameters calibrated with data from the beginning to day 21.

Overall, the model is capable of reproducing and predicting the SWC data accurately and the parameter estimation method is successful with an average calibration relRMSE of 5.85% and average validation relRMSE of 6.8% for the SWC. Calibration is efficient and the optimization problems used to compute a set of parameters are each solved in less than 2 minutes on a modern laptop computer. In particular, the strategy of reusing parameters is successful and for example, the transpiration coefficients calibrated on one data set can be used for all other simulations of a cycle. This allows the estimation of the least amount of parameters possible and results in a faster calibration. Moreover, this demonstrates the generic quality of the model and provides an extra validation of the processes represented.

Calibration was started with the sensors under drippers without lettuces as this required only the calibration of the porosity and the evaporation coefficient. The model is capable of predicting this data with very low errors, with for the first cycle an average calibration relRMSE of 3.2% and an average validation relRMSE of 4.19%. For the second cycle, good results were obtained without recalibration of the porosity parameter, with a mean relRMSE of 7.28%, and this can be explained by the fact that the sensors were not moved and the soil in this area of the bins was left undisturbed between both growth cycles. However, the soil under the lettuces was disturbed after the harvest of the first cycle and sensors were repositioned, leading to a modified local soil structure which affected the porosity. Thus, for the data from sensors under lettuces, the porosity parameter had to be calibrated again for the second cycle.

After calibration of transpiration coefficients and the canopy cover growth rate, the model also simulated accurately SWC data from the sensors under lettuces with, for the first cycle, an average calibration relRMSE of 7.75% and average validation relRMSE of 6.37% and for the second cycle an average calibration relRMSE of 5.47% and average validation relRMSE of 8.14%. For the canopy cover, the model achieves low calibration errors, with a mean relRMSE of 12.3% for the first cycle and of 7.76% for the second cycle. The model is less accurate during the validation period, in particular in the first cycle with an average relRMSE of 49.73% but with a better performance for the second cycle with an average relRMSE of 29.44%. However despite these errors, the model is capable of reproducing the general trend of the canopy growth sufficiently well to obtain very good predictions of the SWC which is the primary objective here.

These results also show that the model is capable of dealing with sensor faults and provides a data filtering method. Indeed, for the first cycle, the SWC data from sensors at position C in the FW bin was of bad quality during the calibration period. Nonetheless, this data was used to estimate the porosity parameters associated with these sensors as well as the canopy growth parameter for the FW water quality. Despite the sensor fault, calibration was successful as shown by the low error obtained during the validation period, for which the data was of better quality. Incidentally, the difference in data quality between the beginning and end periods of cycle 1 in the FW bin likely explains why validation errors are lower than calibration errors for sensors A and C. During the second cycle, sensor D in the TWW bin, recorded an extremely noisy signal but again the calibration method was capable of dealing with this fault and porosity was correctly estimated, yielding low validation errors. Furthermore, in the case of a malfunction resulting in no more recordings, such as sensor A in the FW bin during the second cycle, the model provides a means of estimating the SWC.

Using the calibrated porosity parameters, the VWC data can be converted to PWC and much less spatial variability can be observed in the converted sensor signals and in the simulations. To compare signals of different sensors, the average relative difference is computed. Denoting θ_A and θ_B the signals of 2 sensors and $\bar{\theta}_A$ the mean of a signal, then the relative difference is:

$$\frac{|\theta_A - \theta_B|}{\bar{\theta}_A} \qquad (20)$$

From this, the average is then computed over the time interval of interest. The average of the relative difference between signals from sensors A and C in the TWW bin, cycle 1, is 30.07% in the top and 22.01% in the bottom layer when comparing the signals of VWC. However, once converted to PWC, these differences drop to 5.4% in the top and 5.52% in the bottom. Differences between top and bottom layers are also reduced and in the case of sensors at position C in the TWW bin, cycle 1, the mean difference between the two layers is 11.5% in VWC but only 3.5% in PWC. In addition, the difference between PWC under drippers with or without lettuce is also small at the beginning but becomes more important later in the cycle, when the lettuce has a greater impact on water dynamics.

The porosity parameters estimated are presented in figure 7 and table 3. It can be noticed that the porosity of the bottom layer is always greater or equal than in the top layer. Assuming that differences in porosity are responsible for the spatial variations measured, the approach presented here can also be seen as a method to estimate the spatial variability of the soil water holding capacity. Indeed, the porosity allows to convert PWC to VWC so that the height corresponding to the field capacity of a layer is ϕZS_{fc}. Therefore, since the field capacity is assumed constant when expressed relative to pore volume, the mean and standard deviation of the field capacity is the mean and standard variation of porosity multiplied by ZS_{fc}. In this present study, this gives an average field capacity between 20.74 mm and 25.09 mm with a standard deviation between 3.51 mm and 5.2 mm, which gives valuable information on the spatial variability of soil water holding capacity and is directly usable for irrigation planning.

Table 3 : Statistics of porosity parameters estimated.

	Cycle 1		Cycle 2	
	Mean	Std	Mean	Std
Top	0.352	0.072	0.319	0.054
Bottom	0.386	0.08	0.393	0.057

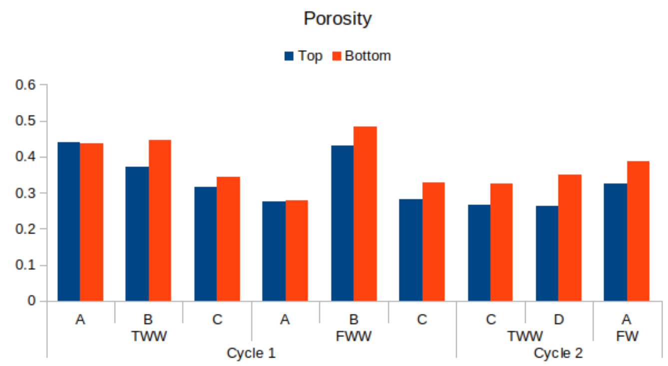

Figure 7 : Porosity parameters calibrated from sensor data. Treated wastewater (TWW) and freshwater (FW) irrigation. A,B,C,D denote sensor positions with A,C,D under a dripper with lettuce and B under a dripper without lettuce.

4.3 Full cycle calibration for analysis

The lettuces biomass were measured only at the end of cycle and therefore to calibrate the biomass growth rate parameter, it was necessary to use data from an entire cycle. Moreover, as the errors on the canopy cover were more important during the second part of both cycles, it was also decided to estimate r_G again in order to reproduce the data more accurately with the objective of using this calibration for the analysis of the experiment. The calibration was successful, with an overall relRMSE of 22.18% for the canopy cover and 2.06% for the biomass (Figure 8).

Figures 9, 10 and 11 compare the data and model simulations of the canopy cover and dry biomass. Differences between bins with irrigation of FW and TWW can be noted during the first cycle mainly, both on the biomass and canopy growth. Interestingly, these differences can be reproduced with the model with differences in only the canopy growth rate whilst the evapotranspiration coefficients and biomass growth rate were the same for both water qualities. The canopy growth rate therefore translates the impact on lettuce growth of the different conditions. Furthermore, the parameters estimated from the two different calibrations can be contrasted to see if growth was different at the beginning and end of the cycles, as presented in table 4. For the first cycle, comparing calibrated parameters with data from the first 21 days or from the full cycle, shows that for the FW bin, lettuces had a slow start with a lower growth rate at the beginning. However, the opposite happened for the TWW bin for cycle 1 and 2 as well as for the FW bin for cycle 2. This can be related to fertilisation of the FW bin occurring only on day 10 of the first cycle, whereas for the TWW, nutrients were added to the soil with each irrigation. Therefore the slower start of the FW lettuces can be explained in part by a nutrient deficit at the beginning of the cycle, with an increased growth rate after fertilisation. This is further supported by the fact that during the second cycle, when both bins were sufficiently fertilised over the entire cycle, lettuces had the similar canopy growth rates for both water qualities.

The different water flows of the soil crop system can be computed from the model simulation and are presented in Figure *12*. During the second cycle, irrigation was done with smaller but more frequent events and the positive impact of this practice can be clearly seen. Indeed, leakage losses were between 43 and 52 % of irrigation in cycle 1 but decreased to between 17 and 28% for cycle 2. Furthermore, despite a decrease in total irrigation of 17mm from the first to the second cycle, the total evapotranspiration increased by 6 mm in the TWW bin and by 15 mm in the FW bin. This is the result of greater evaporation due to higher temperatures and increased transpiration from greater growth of the lettuces during the second cycle.

Table 4 : Canopy growth rate parameters calibrated

	Irrigation water quality	First calibration	Full cycle calibration
Cycle 1	TWW	0.0257	0.0226
	FW	0.01638	0.0184
Cycle 2	TWW	0.05544	0.04631
	FW	0.0497	0.0457

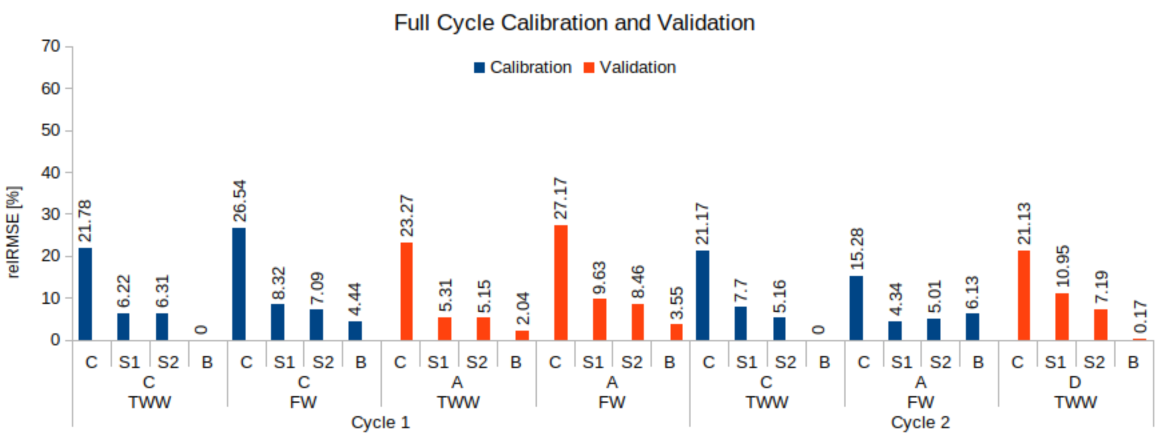

Figure 8 : Calibration and validation errors for 2nd calibration of canopy cover and biomass growth rates with data of entire cycle.Treated wastewater (TWW) and freshwater (FW) irrigation. A,B,C,D denote sensor positions with A,C,D under a dripper with lettuce and B under a dripper without lettuce.

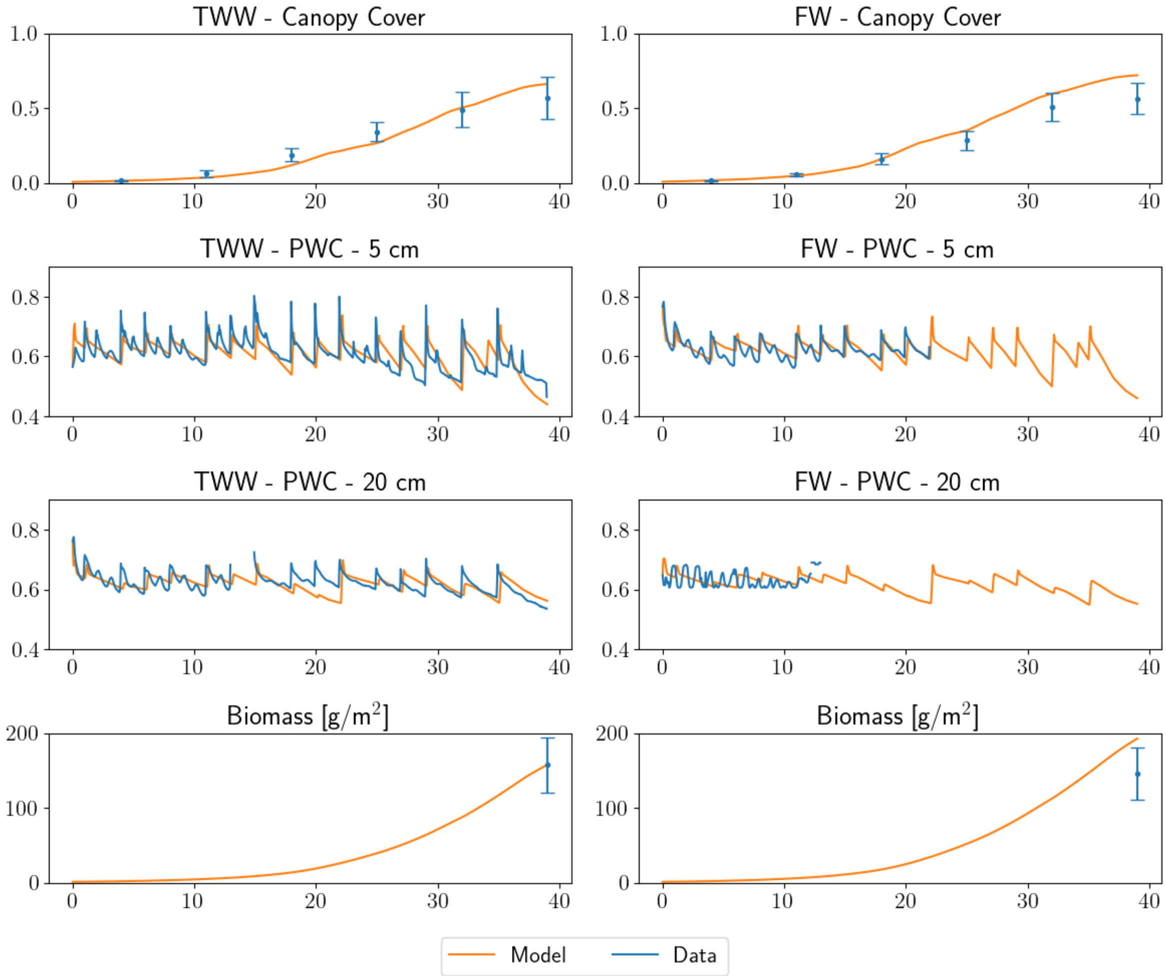

Figure 9 : Comparison of data and model with canopy cover and biomass growth rate parameters calibrated with data from the entire cycle. Left : TWW, cycle 2, sensor position C. Right : FW, cycle 2, sensor position A. Data for soil water is converted to PWC with the model's porosity parameter.

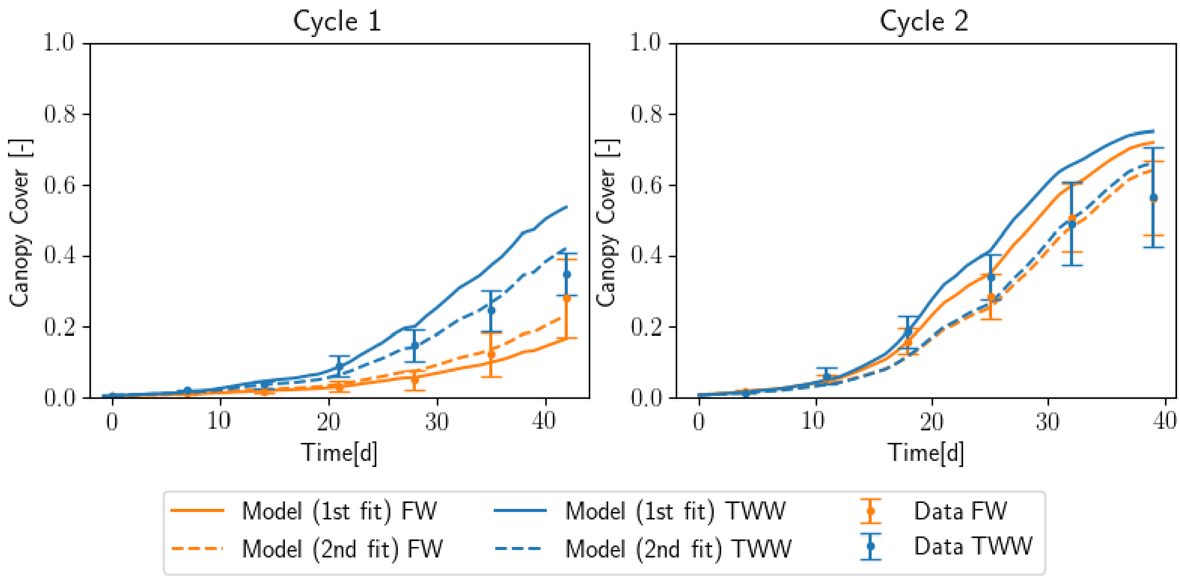

Figure 10 : Comparison of Canopy cover data and model. Data is average per bin with error bars representing one standard deviation. For the model, the first fit was done using only data from the first 3 weeks of each cycle and the 2nd fit was done using data from the entire cycle.

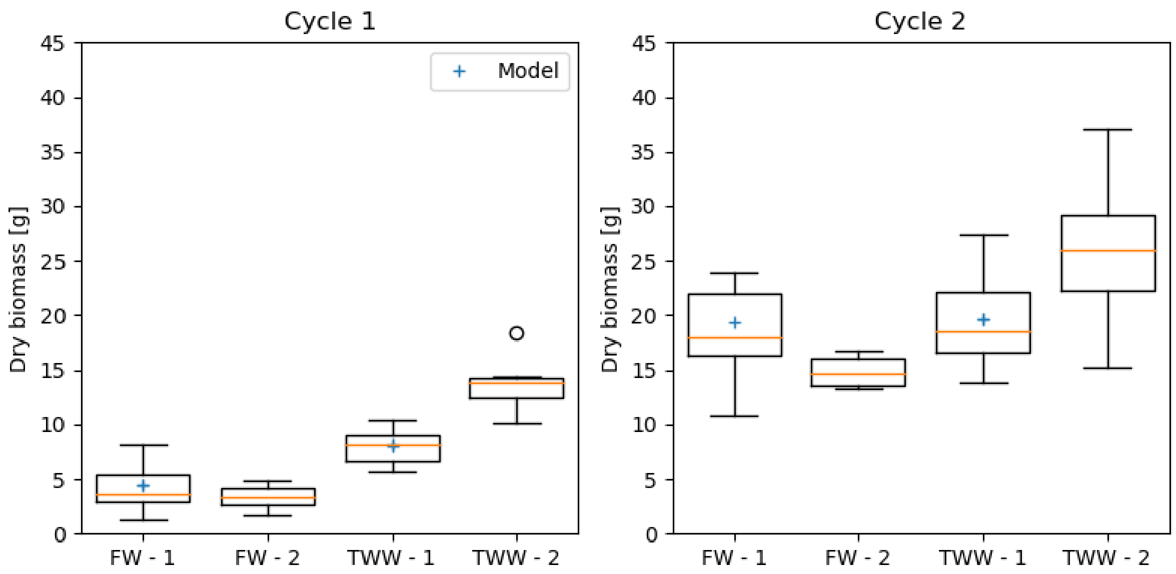

Figure 11 : Dry mass [g] of lettuce at harvest per bin with the soil water sensors positioned in bins labelled 'FW - 1' and 'TWW - 1'. Bins labelled 'FW - 2' and 'TWW - 2' are replicates for which soil moisture sensor data was not available. Simulated final dry biomass is also shown, with the biomass growth parameter calibrated with data from bin 'TWW - 1'.

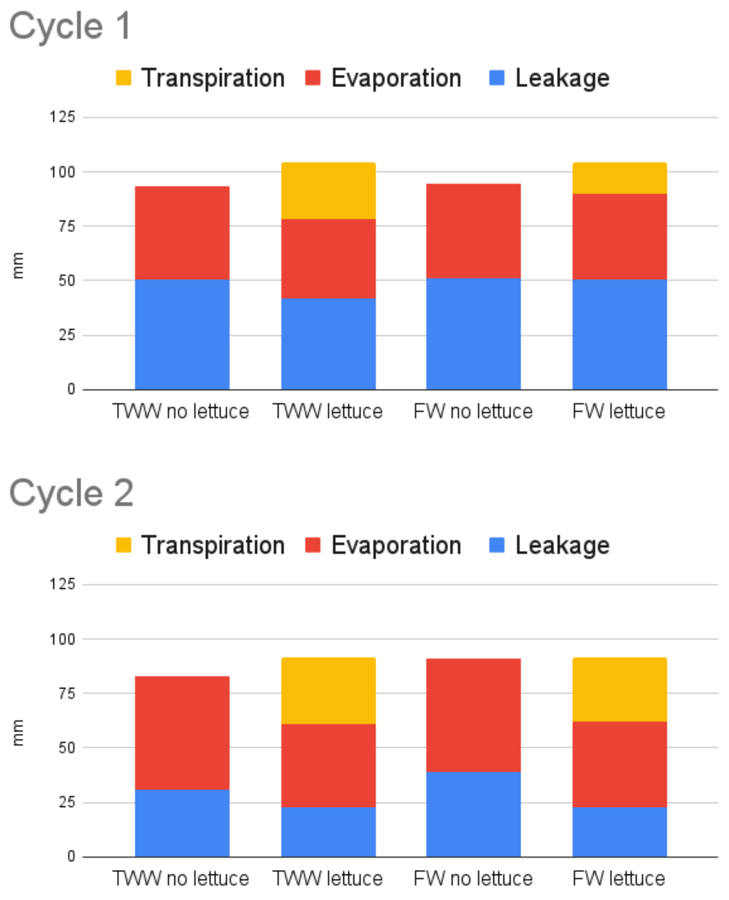

Figure 12 : Water flows computed from model for treated wastewater (TWW) and freshwater (FW) bins.

5 Conclusions

The results of this study show that a simple model can be a valuable tool to complement sensors in accurately estimating the water status of a soil crop system. Indeed, the model and the calibration method are successful at taking into account the spatial variations in the SWC measured, whether they are caused by sensor fault or actual variations in soil water. In particular, this work provides a means to obtain a single value representing the soil water status in the vicinity of a sensor that can be used for irrigation planning.

The simplicity of the model translates into fast calibration and simulations, which are essential for practical use. A disadvantage of a simpler model as the one presented here, is that not all processes involved in crop growth or soil water dynamics are included, leading to a model which can be expected to be less generic. However, trying to take into account all effects to obtain a generic model can lead to extremely complex models which turn out to be very difficult to calibrate efficiently. In general, developing models for real world application implies making trade-offs between model complexity for genericity and simplicity for efficiency. Access to online measurments has an impact on these trade offs, as data can be used to calibrate a simple model to each specific situation, achieving genericity with a simpler representation.

Acknowledgements

These results are part of the JPI project Control4reuse (http://control4reuse.net) financed by the French Research National Agency under the contract ANR-18-IC4W-0002.

References

Abioye, E. A., Abidin, M. S. Z., Mahmud, M. S. A., Buyamin, S., Ishak, M. H. I., Abd Rahman, M. K. I., ... & Ramli, M. S. A. (2020). A review on monitoring and advanced control strategies for precision irrigation. *Computers and Electronics in Agriculture*, *173*, 105441.

Ait-Mouheb, N., Mange, A., Froment, G., Lequette, K., Bru-Adan, V., Maihol, J. C., ... & Wery, N. (2022). Effect of untreated or reclaimed wastewater drip-irrigation for lettuces and leeks on yield, soil and fecal indicators. *Resources, Environment and Sustainability*, *8*, 100053.

Ait-Mouheb, N., Bahri, A., Thayer, B. B., Benyahia, B., Bourrié, G., Cherki, B., ... & Harmand, J. (2018). The reuse of reclaimed water for irrigation around the Mediterranean Rim: a step towards a more virtuous cycle?. *Regional Environmental Change*, *18*(3), 693-705.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). FAO Irrigation and drainage paper No. 56. *Rome: Food and Agriculture Organization of the United Nations*, *56*(97).

Berry D. (2013). Culture biologique des laitues. https://extranet-ain.chambres-agriculture.fr/fileadmin/user_upload/National/FAL_commun/publications/Auvergne-Rhone-Alpes/AB_Culture_biologique_laitues_2013.pdf, (Accesed 18-2-2021)

Bogena, H. R., Huisman, J. A., Oberdörster, C., & Vereecken, H. (2007). Evaluation of a low-cost soil water content sensor for wireless network applications. *Journal of Hydrology*, *344*(1-2), 32-42

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., ... & Sinoquet, H. (2003). An overview of the crop model STICS. *European Journal of agronomy*, *18*(3-4), 309-332.

Cheviron, B., Vervoort, R. W., Albasha, R., Dairon, R., Le Priol, C., & Mailhol, J. C. (2016). A framework to use crop models for multi-objective constrained optimization of irrigation strategies. *Environmental Modelling & Software*, *86*, 145-157.

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water resources research*, *14*(4), 601-604.

Cobbenhagen, A. T. J. R., Antunes, D. J., van de Molengraft, M. J. G., & Heemels, W. P. M. H. (2021). Opportunities for control engineering in arable precision agriculture. *Annual Reviews in Control*, *51*, 47-55.

Cobos, D., & Campbell, C. (2007). Correcting temperature sensitivity of ECH2O soil moisture sensors. *Appl. Note. Decagon Devices, Pullman, WA*.

Feng, G., & Sui, R. (2020). Evaluation and calibration of soil moisture sensors in undisturbed soils. *Transactions of the ASABE*, *63*(2), 265-274.

Khalil, H. (2015) Nonlinear Control, Global Edition, Pearson, 394 pages.

Laio, F., Porporato, A., Ridolfi, L., & Rodriguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: II. Probabilistic soil moisture dynamics. *Advances in water resources*, *24*(7), 707-723.

Mailhol, J. C., Ruelle, P., Walser, S., Schütze, N., & Dejean, C. (2011). Analysis of AET and yield predictions under surface and buried drip irrigation systems using the Crop Model PILOTE and Hydrus-2D. *Agricultural Water Management*, *98*(6), 1033-1044.

Pelak, N., Revelli, R., & Porporato, A. (2017). A dynamical systems framework for crop models: Toward optimal fertilization and irrigation strategies under climatic variability. *Ecological Modelling*, *365*, 80-92.

Pereira, L. S., Paredes, P., & Jovanovic, N. (2020). Soil water balance models for determining crop water and irrigation requirements and irrigation scheduling focusing on the FAO56 method and the dual Kc approach. *Agricultural water management*, *241*, 106357.

Rodríguez-Iturbe, I., & Porporato, A. (2007). *Ecohydrology of water-controlled ecosystems: soil moisture and plant dynamics*. Cambridge University Press.

Šimůnek, J., M. Šejna, and M. Th. van Genuchten, New Features of the Version 3 of the HYDRUS (2D/3D) Computer Software Package, Journal of Hydrology and Hydromechanics, 66(2), 133-142, doi: 10.1515/johh-2017-0050, 2018.

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. *Agronomy Journal*, *101*(3), 426-437.

Stetson, L. E., & Mecham, B. Q. (Eds.). (2011). *Irrigation*. Irrigation Association.

Sui, R., & Vories, E. D. (2020). Comparison of sensor-based and weather-based irrigation scheduling. Applied Engineering in Agriculture, 36(3), 375-386.

Villalobos, F. J., & Fereres, E. (Eds.). (2016). *Principles of agronomy for sustainable agriculture* (p. 443457). New York, NY, USA:: Springer.

Vories, E., & Sudduth, K. (2021). Determining sensor-based field capacity for irrigation scheduling. *Agricultural Water Management*, *250*, 106860.

Walter, E., & Pronzato, L. (1997). *Identification of parametric models: from experimental data*. Springer Verlag.

Warrick, A. W. (2001). *Soil physics companion*. CRC press.