Fine-grained Soil Moisture Monitoring with PLUTO[⋆,⋆⋆]

Matteo Francia[1,†], Joseph Giovanelli[1,*,†] and Matteo Golfarelli[1,†]

[1]DISI – University of Bologna, Via dell'Università 50, 47522 Cesena, Italy

Abstract

Controlling soil moisture is crucial in optimizing watering and crop performance, particularly for crops with high water demands such as Kiwi. Monitoring and simulating soil behavior are two key approaches to understand soil behavior. Proximal sensors are the most reliable way to monitor soil moisture. While in the past sensor costs limited their adoption, the progressive cost reduction makes now possible to properly capture moisture dynamics in the soil volume occupied by roots. Physically-based numerical models can be used to further understand soil moisture dynamics, but solely in an off-line manner due to their time-consuming simulations. We introduce PLUTO, a cost-effective solution that, starting from sensor data, leverages both Physically-based and machine learning models to build on-line moisture profiles for long-term watering optimization. PLUTO, relies on bi/tri dimensional sensor grids that proved to largely overcome the accuracy of previous profiles obtained with traditional sensor layouts. Besides, we provide an analysis of sensor importance that takes in consideration the trade-off between accuracy, number, and position in order to suggest a smart placement.

Keywords

Precision Farming, Smart irrigation, Machine Learning, Kiwi, Sensor Analysis

1. Introduction

Controlling soil moisture is a crucial factor in optimizing watering and crop performance [1]. For instance, Kiwi (*Actinidia deliciosa*), our case study, has high water demand [2] and farmers tend to perform over-watering but this results into fruits with less dry mass, complications in maintenance after harvest, and risks such as groundwater depletion and plant suffocation. There is the need of monitoring soil moisture so that it can be kept at optimal levels, especially in the volume occupied by tree roots. This portion of soil is particularly subject to spatial variability for: (i) uneven root suction, (ii) limited watering-system coverage, (iii) differences

SEBD 2023: 31st Symposium on Advanced Database System, July 02–05, 2023, Galzignano Terme, Padua, Italy

[⋆]Patent pending

[⋆⋆]This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

[*]Corresponding author.

[†]These authors contributed equally.

m.francia@unibo.it (M. Francia); j.giovanelli@unibo.it (J. Giovanelli); matteo.golfarelli@unibo.it (M. Golfarelli)

https://www.unibo.it/sitoweb/m.francia/en (M. Francia); https://www.unibo.it/sitoweb/j.giovanelli/en (J. Giovanelli); https://www.unibo.it/sitoweb/matteo.golfarelli/en (M. Golfarelli)

0000-0002-0805-1051 (M. Francia); 0000-0002-0990-3893 (J. Giovanelli); 0000-0002-0437-0725 (M. Golfarelli)

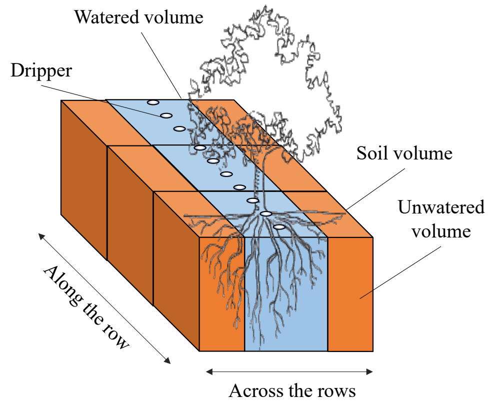


Figure 1: Relevant elements in a orchard.

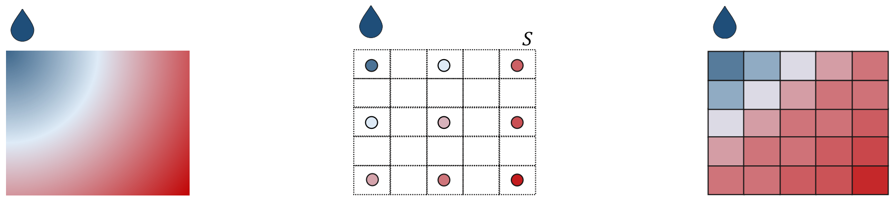
in soil composition, and (iv) exposure to atmospheric agents. The most reliable approach to monitor soil moisture consists of exploiting proximal sensors (i.e., installed below the ground to monitor soil moisture in the plant root zone). Typically, a single (0D) or a line of sensors (1D) are used for cost reasons. This limit the data precision since moisture varies in the soil volume.

Observed moisture data have been historically enhanced with the aid of *Physically-based models* (i.e., numerical solutions that *encode* the physical laws of hydrological fluxes; e.g., HYDRUS [3], CRITERIA [4]). They internally build fine-grained moisture representations, but require (i) long time series of observed moisture data to calibrate the different soil and plant parameters, (ii) an expensive and resource-consuming simulation, and (iii) frequent parameter updates. For these reasons, they have been used to typically carry out *spot* researches on soil moisture dynamics in an off-line manner (e.g., [5], [6], [7], [8], [9]). Recently, *Machine Learning models* (i.e., solutions that learn patterns from a large set of examples; e.g., Artificial Neural Network [10], Support Vector Machine [11]) have been applied. They are acknowledged for being low resource consumer (after an off-line training phase), flexible, and robust. Yet, existing works in literature addressed solely time series forecasting (e.g., [12], [13], [14]) and off-line soil moisture pattern studies (e.g., [15], [16]).

To the best of our knowledge, there is no approach aimed at creating a fine-grained multidimensional representation of soil moisture given a sensor grid. In response, we introduce PLUTO[1] [17]: a cost-effective solution that leverage both Physically-based and Machine Learning models to build on-line moisture profiles, whose reference application is monitoring and *long-term* watering optimization. This is indeed beneficial for preserving optimal soil moisture levels and preventing water waste. We focus on the following contributions.

- The concept of 2D/3D moisture profile as a fine-grained representation of a sensor grid.
- Two alternative solutions to compute the profile, capturing linear and non-linear patterns.
- An analysis of the trade-off between the accuracy, the number, and the position of sensors to smartly place them.

[1]In Greek mythology, god of wealth, linked to the prosperity of crops.



(a) Actual soil moisture. (b) Raw sensor grid (c) Soil profile

Figure 2: Snapshot of soil moisture in a soil slice; the water drop represents a dripper.

Our work is one of the outcomes of the Agro.Big.Data.Science project [18]. The project, funded by Regione Emilia Romagna, aims at studying and implementing digital solutions to support smart and precision farming.

The remainder of this paper is organized as follows. Section 2 provides a formal introduction to the important domain notions for this work. Section 3 illustrates the approach in detail. Section 4 reports the tests carried out on real data collected over two years and a sensor layout analysis for optimal placement Finally, Section 5 draws the conclusions.

2. Domain Formulation

In orchards, where a stable watering system can be built, drip irrigation is widely used as it enables precise watering that, in turn, reduces water waste. Figure 1 shows an example of an orchard watered through a single-pipeline dripper system. The cubes represent the soil volume taken by the roots of each tree. *Along the row* of trees, a limited distance between drippers ensures a homogeneous watered volume (in blue); while *across the rows* (i.e., between two lines of trees), the soil volume remains completely unwatered (in orange).

Our goal is to create a moisture profile that represents the whole soil volume.

Definition 1 (Soil volume)**.** *Given a tree, its* soil volume *is a parallelepiped of soil that contains most of the tree roots. The soil volume is centered in the tree position.*

Definition 2 (Sensor grid)**.** *A* sensor grid $S = \{s^1, ..., s^{|S|}\}$ *is an n-dimensional layout of $|S|$ sensors installed in a soil volume. Each* sensor s^i *is defined by a three-dimensional displacement $(s^i.x_1, s^i.x_2, s^i.x_3)$ with respect to the center of the soil volume, and by a soil moisture value $s^i.v$.*

Depending on n, the grid resembles a line ($n = 1$), a rectangle ($n = 2$) or a parallelepiped ($n = 3$). The monitored value depends on the sensor technology, typically sensors measure volumetric water content (i.e., the volume of liquid water per volume of soil) or the soil potential (i.e., the energy required by tree roots to extract water from soil particles).

Definition 3 (Moisture profile)**.** *Given an n-dimensional sensor grid S, the* moisture profile *is an n-dimensional grid $P = \{p^1, ..., p^{|P|}\}$ that approximates, in each p^i, the soil moisture measured by S. P is* fine-grained *with respect to S since $|P| > |S|$.*

Building the moisture profile, based on raw sensor measurements, requires to estimate soil moisture of the whole soil volume at a fine-grained resolution.

Example 1. *Given a 2D sensor grid covering an area of* $0.6m^2$ *(i.e., a rectangle with a width of* $1m$ *and a height of* $0.6m$*), we employed a sensor grid with 12 sensors (i.e.,* $|S| = 12$*) and a moisture profile with 1000 points (i.e.,* $|P| = 1000$*). As a result, we obtained a moisture profile having a granularity of* $6cm^2$ *(i.e.,* $0.6m^2/1000$*) while the sensor grid granularity is* $500cm^2$ *(i.e.,* $0.6m^2/12$*). For the sake of clarity, Figure 2b shows a sensor grid with* $|S| = 9$ *and Figure 2c shows a moisture profile with* $|P| = 25$.

3. PLUTO

The transformation of raw sensor measurements into a fine-grained moisture profile is achieved through a profiling function.

Definition 4 (Profiling function)**.** *Given an* n*-dimensional sensor grid* S *and a moisture profile* P*, a* profiling function $f : S \to P$ *approximates the moisture profile* P *starting from the grid* S.

The role of a profiling function is approximating the soil moisture values in those positions of the moisture profile where a sensor is not present. A profiling function is based on sensor grid measurements and can optionally exploit further information about the behavior of the soil. Several profiling functions can be adopted. We propose two alternative approaches that differ in the information exploited.

- *Soil-feature unaware* - FU: exploits the sensor measurements only. The most obvious choice is to carry out a linear interpolation between pairs of sensor values.
- *Soil-feature aware* - FA: exploits the knowledge about soil hydrological dynamics to keep into account non-linearities and to get a more accurate estimation.

Sample profiles created through the two profiling functions are reported in Figure 3. It is apparent that the FA one encodes complex non-linearities deriving from the combined effect of the multiple factors that characterize the soil and the plant. PLUTO data processes for the two alternative profiling functions are sketched in Figure 4. The FU function does not require to be fitted to a specific field and only gets sensor data as input, while the FA function is trained offline to capture non-linear behaviors due to specific plant and soil behaviors at hand.

As to the FU profiling function, we rely on the well-known n-linear interpolation, where n is the profile grid dimensionality. Although this regression approach is bounded to linear behavior between sensor pairs, we highlight that the composition of several linear strokes approximate a non-linear trend. For the sake of conciseness, in the following, we describe the 2-linear case. Given a 2D sensor regular grid S, this technique carries out a linear interpolation in each dimension independently from each other. The approach consists of two phases (Figure 5). For each point $p \in P$ of the moisture profile to be computed: (i) we find the four sensors S that determine the minimum bounding rectangle enclosing p (Figure 5a), then (ii) we compute $p.v$ (Figure 5b) by interpolating along the x_1 axis first. Then, exploiting the obtained points r (blue dots), interpolation is performed along the x_2 axis and the value $p.v$ is finally determined. The trilinear procedure is analogous: it just has three steps instead of two.

As to the FA function, we rely on a machine learning model to capture non-linear soil moisture behaviors; in particular we built an Artificial Neural Network (ANN) that learns the moisture

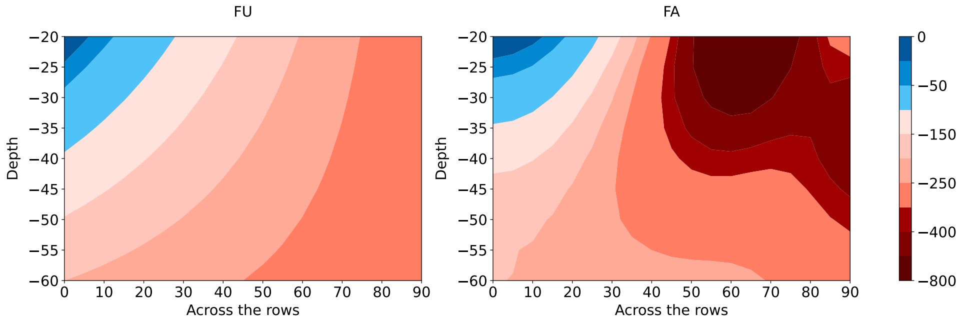


Figure 3: Moisture profile built by the two types of profiling functions, namely feature unaware (left) and feature aware (right).

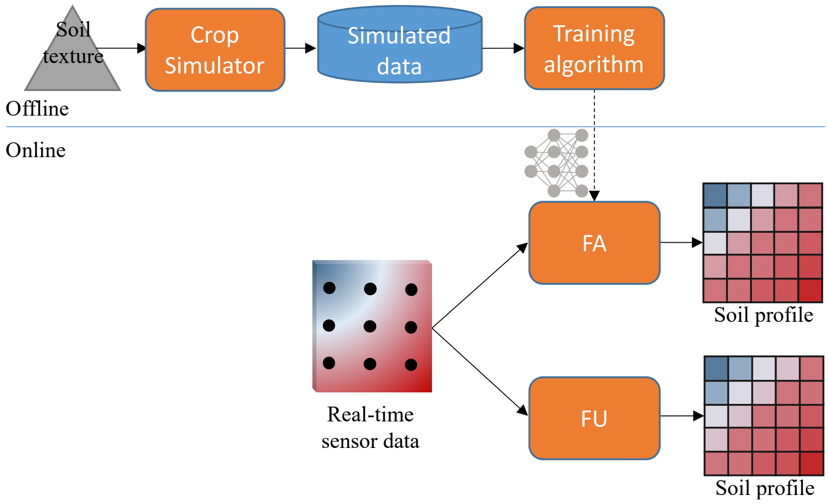
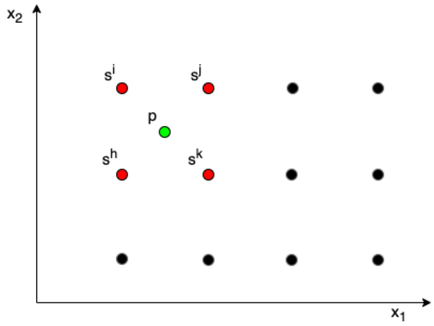
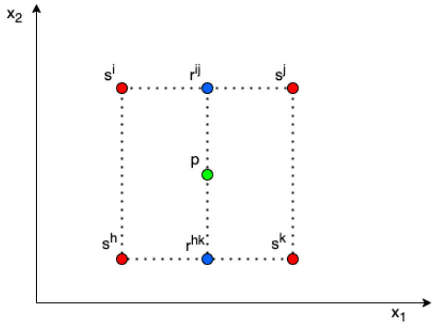


Figure 4: Feature-aware and feature-unaware processes for building a moisture profile.

dynamics from a physically-based model (i.e., CRITERIA [4]). Once trained the ANN, given the sensor measurements, efficiently computes the moisture values at the profile granularity. This would be not possible directly using the numerical model that requires higher resources. The learning process is sketched in Figure 4. The farmer provides the main characteristics of the implant (e.g., soil texture, sensor layout) so that the crop simulator is set up accordingly to reproduce the hydrological fluxes in the case at hand. Then, historical weather conditions and watering sessions (a period of 4-month period has proved to be sufficient) are fed to CRITERIA to generate the training data for the ANN: the input layer has one neuron for each sensor to interpolate and the output layer has as many neurons as the number of points in P. Figure 6 and Table 1 show the ANN architecture and the related hyper-parameters tuned with the aid of a hyper-parameter tuning process implemented with HyperOpt [19]. HyperOpt exploits state-of-the-art optimization techniques to heuristically explore search spaces of hyper-parameters. The ANN learns the soil behavior from simulated data in an off-line phase, which is carried out only once at the time of installation of the system.



(a) Bounding rectangle for p



(b) Bilinear interpolation

Figure 5: A 2D example of the feature unaware function.

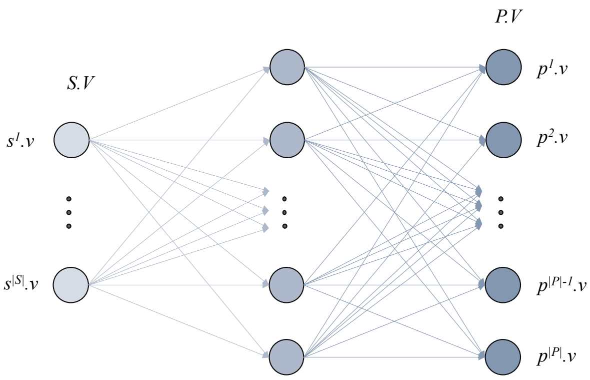


Figure 6
ANN layout.

Hyper-parameter	Value
# Hidden Layers	1
# Neurons per layer	100
Activation function	Tanh
Normalization	Z-score
# Training epochs	50
Batch size	30
Reduce learning rate	factor=2, patient=10

Table 1
ANN hyper-parameters.

4. Empirical Evaluation and Layout Analysis

We tested PLUTO one of the Agro.Big.Data.Science project installations [18] located in Faenza (Emilia Romagna, Italy). This orchard is watered through a single pipeline of drippers (distance between drippers 40cm) and Kiwifruit vines were spaced 2m along the row and 4.5m across the rows. Soil moisture is monitored through two installations, specifically 2D and 3D sensor grids of – respectivelly – 12 and 15 gypsum block sensors. We monitored the soil water potential, hence the profile values are in unit of pressure (namely, in cbar).

In real-world applications, the number of sensors is by far less than those we used in our research project. Section 4.1 leverages an incremental number of sensors to compute the profile and the remaining as ground truth to evaluate the performance. Section 4.2 conducts an analysis of the sensor importance according to the trade-off between accuracy and position.

4.1. Performance Evaluation

Figure 7 and Figure 9 show the system performance for – respectivelly – the 2D and 3D cases, varying the number of input sensors. In each test, the performance is calculated as the RMSE

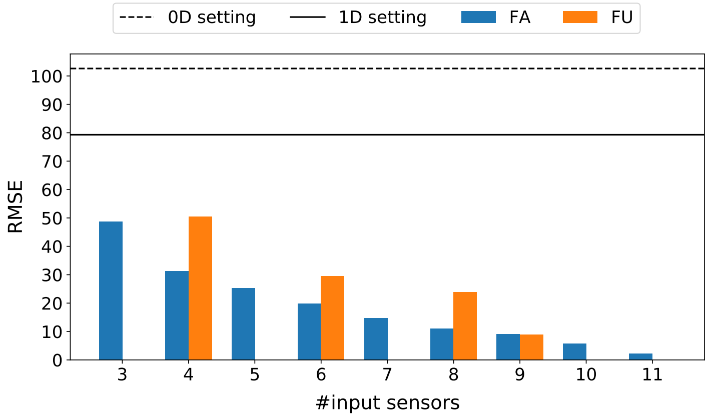


Figure 7
2D profile performances as a function of the input sensors. Single sensor (0D) and column (1D) settings are reported for comparison.

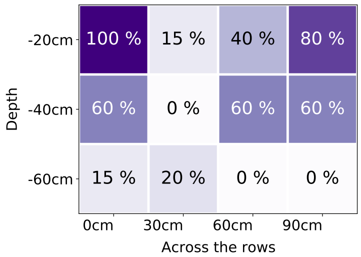


Figure 8
Frequency of times each sensor appeared in the best layouts.

(i.e., Root Means Square Error, the less the better) between the ground truth and the estimated values by the profile. To better understand the advantage provided by PLUTO, we also report the performance of profiling based on a single sensor — 0D setting — (i.e., the *de facto* standard) or a column of 3 sensors at different depths — 1D setting.

To extend the soil moisture value of the single sensor to the entire soil volume, we must assume the soil moisture to be constant in the whole volume. Similarly, when a column of sensors is available, we assume that soil moisture is constant at the same depth in the soil. Since the accuracy varies based on the sensor location, we choose the single sensor position (or the sensor line) that minimizes the RMSE as a fair comparison baseline. The same methodology was adopted to displace the sensors used to calculate the profile with our approach. Specifically, 2D profiles assume soil moisture to be constant along the row, while 3D profiles have no assump

In both cases, the 0D setting estimations are by far less accurate than the ones obtained by our profiling functions. The 1D settings achieve slightly better results since it captures soil moisture at different soil depths, but fails in capturing longitudinal variations. The extent of these errors is not negligible since the optimal range for soil moisture for kiwi cultivation is $[-100; -300)$ cbar [20].

RMSE for FA and FU gradually decreases as the number of sensors increases. Note that the linear profiling function can be computed only for some sensor layouts, due to the intrinsic geometrical constraints (i.e., the profile region must be partitioned in bounding rectangles/cubes). The ANN profiling function always outperforms the bilinear one due to its capability to better model non-linear behaviors.

4.2. Sensor Layout Analysis

As shown in Figures 7 and 9, accuracy varies with the number of sensors. Given a regular grid with n sensors, several layouts are possible when $m < n$ sensors are used. In particular, while the bilinear FU function implies some geometric constraints, in the ANN-based FA one all the

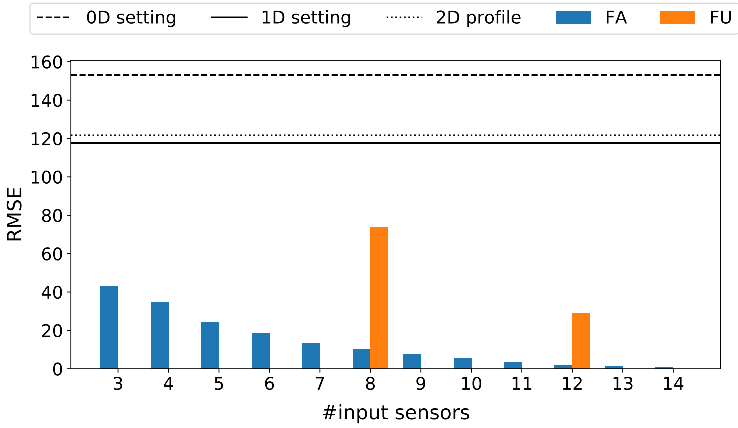


Figure 9: 3D profile performances as a function of the input sensors. Single sensor (0D), column (1D) settings, and (2D) profile are reported for comparison.

layouts are feasible. To compare different layout performances and suggest the smart one, we considered the five profiles that achieve the best performances when $m \in [3, 6]$ sensors are used. Figure 8 shows the percentage of times the grid sensor appears in one of the top-performing layouts. Noticeably, the layouts with highest performance include the sensors that convey more information on soil moisture. In particular: (i) the sensor just under the dripper $(0cm, -20cm)$ since it is the most affected by the effects of the dripper; (ii) the sensor near the surface farthest from the dripper $(90cm, -20cm)$ since it records the state of the unwatered volume and is strongly influenced by atmospheric phenomena (e.g., sun, rain); (iii) one sensor at mid-depth $(*, -40cm)$ since it captures the soil behavior when not directly affected by watering and atmospheric phenomena.

5. Conclusions and Future Work

We presented PLUTO [17], an original approach to compute fine-grained moisture profiles. PLUTO relies on a grid of soil moisture sensors and it largely outperforms previous approaches based on a single or a column of sensors. We have shown that three sensors, properly placed in the soil, are sufficient to effectively obtain the profile.

We are currently turning our monitoring system into a forecasting one. We are testing ANNs to create a solution that initially (i.e., before the deployment of the sensors) learns from a soil simulator and then improves its accuracy by exploiting real sensor samples collected during operations. The overall goal is to create a prescriptive analytics system that automatically activates the watering system based on a soil moisture prediction module fed with weather forecasts.

References

[1] T. Turkeltaub, D. Kurtzman, O. Dahan, Real-time monitoring of nitrate transport in the deep vadose zone under a crop field–implications for groundwater protection, Hydrology and Earth System Sciences 20 (2016) 3099–3108.

[2] M. Judd, K. McAneney, M. Trought, Water use by sheltered kiwifruit under advective conditions, New Zealand journal of agricultural research 29 (1986) 83–92.

[3] J. Šimůnek, M. Van Genuchten, M. Šejna, Development and applications of the hydrus and stanmod software packages and related codes, Vadose Zone Journal 7 (2008) 587–600. doi:10.2136/vzj2007.0077.

[4] M. Bittelli, A. Pistocchi, F. Tomei, P. Roggero, R. Orsini, M. Toderi, G. Antolini, M. Flury, CRITERIA-3D: A mechanistic model for surface and subsurface hydrology for small catchments, Cabi, 2011.

[5] Z. Pan, Y. Tong, J. Hou, J. Zheng, Y. Kang, Y. Wang, C. Cao, Hole irrigation process simulation using a soil water dynamical model with parameter inversion method, Agricultural Water Management 245 (2021). doi:10.1016/j.agwat.2020.106542.

[6] H. Li, J. Yi, J. Zhang, Y. Zhao, B. Si, R. Hill, L. Cui, X. Liu, Modeling of soil water and salt dynamics and its effects on root water uptake in heihe arid wetland, gansu, china, Water (Switzerland) 7 (2015) 2382–2401. doi:10.3390/w7052382.

[7] G. Egea, A. Diaz-Espejo, J. Fernández, Soil moisture dynamics in a hedgerow olive orchard underwell-watered and deficit irrigation regimes: Assessment, predictionand scenario analysis, Agricultural Water Management 164 (2016) 197–211. doi:10.1016/j.agwat.2015.10.034.

[8] M. Cordeiro, V. Krahn, R. Ranjan, S. Sager, Water table contribution and diurnal water redistribution within the corn root zone, Canadian Biosystems Engineering / Le Genie des biosystems au Canada 57 (2016) 139–148. doi:10.7451/CBE.2015.57.1.39.

[9] A. Zapata-Sierra, J. Roldán-Cañas, R. Reyes-Requena, M. Moreno-Pérez, Study of the wet bulb in stratified soils (sand-covered soil) in intensive greenhouse agriculture under drip irrigation by calibrating the hydrus-3d model, Water (Switzerland) 13 (2021). doi:10.3390/w13050600.

[10] A. Krogh, What are artificial neural networks?, Nature biotechnology 26 (2008) 195–197.

[11] W. S. Noble, What is a support vector machine?, Nature biotechnology 24 (2006) 1565–1567.

[12] A. Goldstein, L. Fink, A. Meitin, S. Bohadana, O. Lutenberg, G. Ravid, Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge, Precision Agriculture 19 (2018) 421–444. doi:10.1007/s11119-017-9527-4.

[13] A. Jiménez, B. Ortiz, L. Bondesan, G. Morata, D. Damianidis, Evaluation of two recurrent neural network methods for prediction of irrigation rate and timing, Transactions of the ASABE 63 (2020) 1327–1348. doi:10.13031/TRANS.13765.

[14] Z. Liang, X. Liu, T. Zou, J. Xiao, Adaptive prediction of water droplet infiltration effectiveness of sprinkler irrigation using regularized sparse autoencoder–adaptive network-based fuzzy inference system (rsae–anfis), Water (Switzerland) 13 (2021). doi:10.3390/w13060791.

[15] C. Arif, M. Mizoguchi, B. I. Setiawan, et al., Estimation of soil moisture in paddy field

using artificial neural networks, arXiv preprint arXiv:1303.1868 (2013).

[16] E. Babaeian, S. Paheding, N. Siddique, V. Devabhaktuni, M. Tuller, Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning, Remote Sensing of Environment 260 (2021). doi:10.1016/j.rse.2021.112434.

[17] M. Francia, J. Giovanelli, M. Golfarelli, Multi-sensor profiling for precision soil-moisture monitoring, Computers and Electronics in Agriculture 197 (2022) 106924.

[18] A.B.D.S., The Agro.Big.Data.Science project, http://agrobigdatascience.it/, 2021. Last accessed: 2021-10-18.

[19] J. Bergstra, B. Komer, C. Eliasmith, D. Yamins, D. D. Cox, Hyperopt: a python library for model selection and hyperparameter optimization, Computational Science & Discovery 8 (2015) 014008.

[20] S. Miller, G. Smith, H. Boldingh, A. Johansson, Effects of water stress on fruit quality attributes of kiwifruit, Annals of botany 81 (1998) 73–81.