



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE  
DELLA RICERCA

## Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Multi-sensor profiling for precision soil-moisture monitoring

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Francia M., Giovanelli J., Golfarelli M. (2022). Multi-sensor profiling for precision soil-moisture monitoring. *COMPUTERS AND ELECTRONICS IN AGRICULTURE*, 197, 1-11 [10.1016/j.compag.2022.106924].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/885669> since: 2022-05-17

*Published:*

DOI: <http://doi.org/10.1016/j.compag.2022.106924>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

# Multi-Sensor Profiling for Precision Soil-Moisture Monitoring<sup>☆,☆☆</sup>

Matteo Francia, Joseph Giovanelli, Matteo Golfarelli\*

*DISI – University of Bologna, Via dell’Università 50, 47522 Cesena, Italy*

---

## Abstract

Controlling soil moisture is crucial in optimizing watering and crop performance. Traditional monitoring systems rely on a single sensor or on a column of sensors that do not allow farmers to properly capture soil moisture dynamics in the soil volume occupied by roots. In this paper we propose PLUTO, an original approach that builds fine-grained 2D and 3D soil moisture profiles by relying on a grid of sensors. Profiles are computed using both interpolation-based and machine learning approaches. Besides the technical description of the approach, the paper reports a set of original visualizations and a large set of tests computed, over two years, on real Kiwi orchards. PLUTO proved to largely overcome the accuracy of profiles obtained with traditional sensor layouts. Considering that the cost of sensors is progressively decreasing, PLUTO provides a cost-effective, operative, and precise solution to moisture monitoring.

---

<sup>☆</sup>Patent pending

<sup>☆☆</sup>This work is partially supported by the AGRO.BIG.DATA.SCIENCE project grant type POR FESR Program 2014-2020, 1-1.2.2.

\*Corresponding author

*Email addresses:* [m.francia@unibo.it](mailto:m.francia@unibo.it) (Matteo Francia), [j.giovanelli@unibo.it](mailto:j.giovanelli@unibo.it) (Joseph Giovanelli), [matteo.golfarelli@unibo.it](mailto:matteo.golfarelli@unibo.it) (Matteo Golfarelli)

*Keywords:* Smart watering, Smart irrigation, IoT, Kiwi, Artificial Neural Network

---

## 1. Introduction

Controlling soil moisture is a crucial factor in optimizing watering and crop performance [1]. For example, Kiwi (*Actinidia deliciosa*), our case study, has high water demand [2]. Indeed, Kiwi is an irrigated crop in most of the areas where it is cultivated such as New Zealand, Chile, and Italy.

Different types of watering systems may be adopted depending on the farming features and needs; typical solutions exploit drippers, sprinklers, hose reels, and flooding. 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.

The main application of our work is preserving an optimal moisture level without wasting water. The risks of over-watering range from groundwater depletion to plant suffocation. As to kiwi, farmers tend to over-water since it leads to larger fruits, but this reduces their dry mass and jeopardizes their maintenance after harvest. Ideally, the moisture level should be known and optimal on the whole soil volume taken by the tree roots. Such volume is subject to strong horizontal and vertical variability caused by many related factors, among them: (i) uneven root suction, (ii) limited watering-system coverage, and (iii) difference in the soil layers in terms of composition and exposure to atmospheric agents.

Figure 1 shows an example of an orchard watered through a single-pipeline dripper system. Each cube shows the soil volume taken by the roots. On

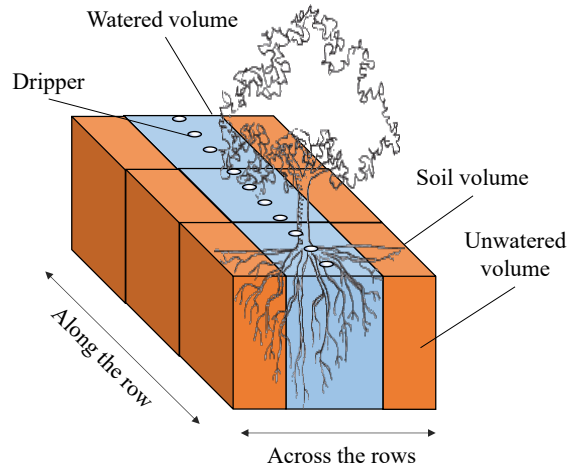


Figure 1: Relevant elements in a orchard.

the one hand, the limited distance between drippers ensures homogeneous moisture (i.e., the watered volume in blue) *along the row* (i.e., multiple trees in the same line). On the other hand, the effect of watering drops *across the rows* (i.e., between two lines of trees), and a portion of the soil volume remains completely unwatered (i.e., the volume in orange).

When the watered volume is symmetric along the row, as in Figure 1, a 2D grid (Figure 2-left) is sufficient to represent the entire soil volume. Conversely, when relevant moisture variations take place along the row too, a 3D grid (Figure 2-right) is more suitable. Such variations may be determined, for example, by too sparse drippers or by non-homogeneous suction of the roots.

Our approach, named PLUTO<sup>1</sup>, focuses on the following contributions.

- The concept of 2D/3D moisture profile (i.e., a fine-grained representa-

---

<sup>1</sup>In Greek mythology, god of wealth. His name is linked to the prosperity of crops.

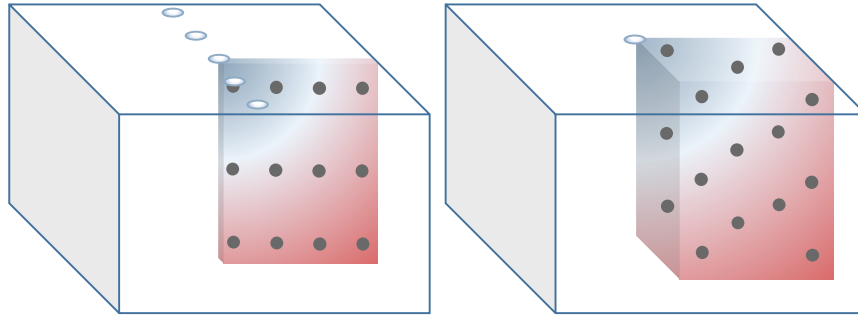


Figure 2: Regular 2D (left) and 3D (right) sensor grids.

tion of moisture in a portion of soil) based on a grid of sensors. The grain of the profile is in the order of magnitude of square/cubic centimeters.

- Two alternative solutions to estimate a fine-grained profile by adopting linear and non-linear approximations. In particular, non-linear approximation relies on a neural network that learns a model of soil moisture given the soil texture.
- An analysis of the trade-off between the accuracy, the number, and the position of sensors.
- Several original profile visualizations that enable visual exploitation of the profile.

To the best of our knowledge, PLUTO is the only approach aimed at creating a fine-grained multidimensional profile of soil moisture given a sensor grid. In particular, previous approaches building a soil profile exploit sensors to calibrate a simulator (e.g., HYDRUS [3]) and thus (i) require long time series of data from the field; (ii) bind the profile exploitation to an expensive

and resource-consuming simulation; and (iii) require frequent updates. For these reasons, such approaches typically carry out *spot* researches to study soil moisture dynamics. Conversely, our goal is to provide a cost-effective, operative solution that monitors soil moisture and whose reference application is *long-term* watering optimization.

The remainder of this paper is organized as follows. Section 2 introduces the basics of soil moisture monitoring and sensors. Section 3 surveys the previous approaches in the area of moisture monitoring. Section 4 describes PLUTO in detail. Finally, Section 5 reports the tests carried out on real data collected over two years, and Section 6 draws the conclusions.

We emphasize that our work is one of the outcomes of the Agro.Big.Data.Science project [4]. The project, funded by Regione Emilia Romagna, aims at studying and implementing digital solutions to support smart and precision farming.

## 2. Background

The monitoring of soil moisture has been proven beneficial in many various applications. For instance, Bordoni et al. [5] estimate the soil moisture on a test slope of Oltrepò Pavese (northern Italy) for the assessment of shallow landslides triggering. Besides, agriculture has witnessed different analyses based on the adopted watering techniques. Among them, the investigation of the water demand of a mature citrus orchard to optimize the water budget during rainy and dry seasons in southern China [6]. In the field of flood irrigation, Hamilton et al. [7] collect soil moisture measurements and compare three contrasting soil management treatments, discovering how to improve

efficiency. In the following, we consider only papers related to precision farming, where it is important to have measurements as fine-grained as possible.

The most common practice is to observe the change in soil moisture with the aid of sensors. According to the displacement, sensors can be categorized as *proximal* or *remote*. Proximal sensors are typically installed below the ground to monitor soil moisture in the plant root zone, offering precise measurements. Conversely, remote sensors (e.g., Unmanned Aerial Systems) have been used for the discovery of global soil moisture patterns, due to their coarse spatial resolutions. Babaeian et al. [8] integrate both technologies to conduct analyses on global soil moisture patterns as well as near the plant. Yet, commercial precision farming installations encompass only the use of proximal sensors, as including both sensor types is not worth the cost. Furthermore, depending on the operating principle, proximal sensors measure either the *volumetric water content* (i.e., the volume of liquid water per volume of soil) or the *soil water potential* (i.e., the energy required by tree roots to extract water from soil particles).

Proximal sensors installations can be categorized into *single-sensor* and *multi-sensor* installations. As the name suggests, the single sensor installation (0D) provides a punctual measurement of soil moisture. Although this installation [9] is cheap, it cannot monitor the soil moisture at different depths. To do so, multi-sensor installations are devised (e.g., ground-based proximal sensors are organized so that a grid is formed). According to the layout, we distinguish: mono-dimensional (1D), bi-dimensional (2D), and three-dimensional (3D) grids. The most popular configuration is the mono-dimensional grid [10, 11, 12, 13, 14], where the sensors are vertically aligned at

different depths. Bi-dimensional [15, 16] and three-dimensional [17, 18] grids are quite less used due to the higher required number of sensors. Yet, they allow monitoring soil moisture changes not only at different depths but also along the other axes. Bi-dimensional grids involve installations of multiple sensors per soil layer, forming thus a matrix that enables the understanding of patterns in an entire slice of soil. While, three-dimensional grids can be seen as more bi-dimensional grids side by side, allowing the examination of more than one slice. An exhaustive dissertation about sensors and data collection can be found in [19].

### 3. Related Literature

Raw sensory data are only the starting point for extracting richer representations of soil moisture. We review the literature of approaches that have been proposed to create such representations. We distinguish the following techniques.

- *Physically-based numerical models* (e.g., HYDRUS [3], CRITERIA [20]) are analytical models that *code* the physical laws determining the hydrological dynamics.
- *Machine learning models* leverage artificial intelligence techniques to *learn* hydrological dynamics from a large set of examples.

Also, we distinguish the following goals.

- *Simulation* estimates the soil dynamics over time with the aim of understanding the soil behavior under specific circumstances. Simulation



is not bound to actual weather and soil moisture values, but rather tests different scenarios to understand the overall soil behavior.

- *Forecasting*, starting from the current moisture values in a specific field, estimates its future values in order to let the user take appropriate decisions (e.g., watering).
- *Profiling* estimates a fine-grained soil representation combining a coarse description of the moisture in the soil with statistical assumptions and the knowledge of soil characteristics.

Although the boundaries between these three goals are blurred, we can easily differentiate them considering that, while forecasting and simulation produce estimations of *future* moisture values, profiling produces a more detailed estimate of *present* values.

Physically-based numerical models have been for a long time the only choice to study hydrological dynamics, and thus are considered well-accepted and solid tools. In this regard, several recent works exploit them with the aim of simulation: Pan et al. [13] investigate the differences between two hole watering methods; Li et al. [14] study the effects of soil water and salt dynamics on root water uptake; Egea et al. [15] analyze the suitability of the irrigation management of a hedgerow olive orchard in different soil types; Cordeiro et al. [16] determine the water table contribution in a cornfield; Zapata-Sierra et al. [17] monitor the evolution of soil moisture in a pepper crop field. As to works that focus on forecasting: Chen et al. [21] predict the soil wetness status in several sites close to catchments of significant size (soil type and leaf area index were the key parameters affecting the model

performance); Shein et al. [22] validate the efficiency of HYDRUS-1D for predicting soil moisture and temperature dynamics with hysteresis in clay loam soils.

Machine learning models in the field of precision agriculture are relatively new in comparison to physically-based numerical models. Karandish et al. [10] evaluate and compare the goodness of machine learning models with physically-based numerical models in several soil moisture simulations. Results assess the efficacy of the former, which additionally benefit of some advantages, above all low resources consumption: no on-line heavy computation is required and the result is rapidly available. As a matter of fact, several approaches apply machine learning models with the aim of simulation: Arif et al. [9] build a model to monitor soil moisture in paddy field with limited meteorological data; Babaeian et al. [8] develop a machine learning model that analyzes surface, near-surface, and root zone soil moisture by exploiting the fusion of remote and proximal sensors. As to forecasting: Liang et al. [18] quantify the water droplet infiltration in a sprinkler irrigation system to improve and facilitate the watering management, Jiménez et al. [12] and Goldstein et al. [11] directly predict the watering volumes recommended by the agronomist.

Finally, there are approaches that leverage physically-based numerical models in synergy with machine learning. For instance, Hinnell et al. [23] exploit HYDRUS as a physically-based numerical model to provide fine-grained samples and Artificial Neural Networks as a machine learning model to extract subsurface wetting patterns. With respect to PLUTO, Hinnell et al. (i) produce a statistical representation of the patterns that is based on a

Table 1: Classification and comparison of related works in the field of precision farming.

Reference	Sensor Type	Layout	Goal	Approach
[13]	Volumetric	multi (1D)	Simulation	PbN
[14]	Volumetric	multi (1D)	Simulation	PbN
[15]	Volumetric	multi (2D)	Simulation	PbN
[16]	Volumetric	multi (2D)	Simulation	PbN
[17]	Volumetric	multi (3D)	Simulation	PbN
[21]	Volumetric	multi (1D)	Forecasting	PbN
[22]	Potential	multi (1D)	Forecasting	PbN
[9]	Volumetric	single (0D)	Simulation	ML
[8]	UAS, Volumetric	multi (1D)	Simulation	ML
[11]	Volumetric	multi (1D)	Forecasting	ML
[12]	Potential	multi (1D)	Forecasting	ML
[18]	Volumetric	multi (3D)	Forecasting	ML
[23]	-	-	Profiling	PbN, ML
PLUTO	Potential	multi (2D/3D)	Profiling	PbN, ML

Unmanned Aerial Systems (UAS), Physically-based Numerical (PbN), Machine Learning (ML).

*single* dripper and assume a *uniform* characterization of the soil, (ii) derive the wetting patterns from generic soil parameters and not from actual sensor values.

Table 1 summarizes the related works described so far and compares their features against PLUTO. To the best of our knowledge, PLUTO is the only one that builds a fine-grained 2D/3D moisture profile based on a coarser sensors grid, through linear and non-linear techniques.

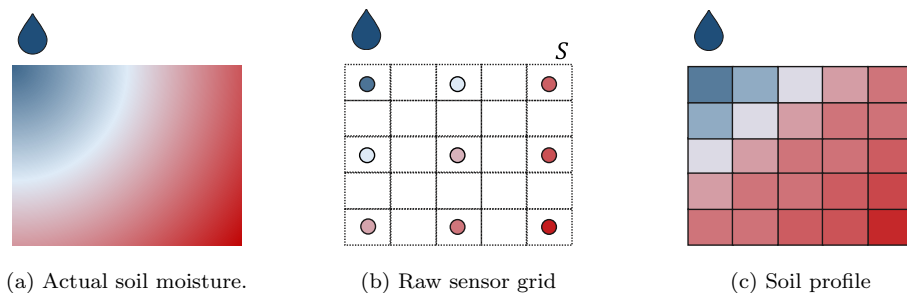


Figure 3: Snapshot of soil moisture in a soil slice; the water drop represents a dripper.

#### 4. Materials and Methods: the PLUTO Approach

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  $S = \{s^1, \dots, s^{|S|}\}$  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 or the soil potential. Figure 4 shows one of the Agro.Big.Data.Science project installations [4] located in Faenza (Emilia Romagna, Italy). This orchard is watered through a single pipeline of drippers (distance between drippers 40cm) and



Figure 4: A sensor grid is monitoring soil moisture close to a kiwi tree near Faenza, Italy. The watering system is composed of single-pipeline drippers.

soil moisture is monitored through a 2D sensor grid of 12 gypsum block sensors.

Soil moisture varies continuously within the soil volume while the raw sensors provide point-wise measurements. Furthermore, in most real-world applications, the number of sensors is by far less than those we used in our research project; this further reduces the comprehensiveness of the measurements. For instance, companies involved in Agro.Big.Data.Science [4] relied on 1 to 3 sensors disposed at different depths (i.e., a 0D or a 1D grid). The moisture profile, based on raw sensor measurements, estimates soil moisture of the whole soil volume at a fine-grained resolution.

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

The approximation  $p^i.v$  is assumed to be constant in the region surrounding  $p^i$ , whose granularity depends on  $|P|$ .

**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 3b shows a sensor grid with  $|S| = 9$  and Figure 3c shows a moisture profile with  $|P| = 25$ .*

A moisture profile covers the sensor grid region that is typically smaller, in size and dimensionality, than the soil volume. To provide a comprehensive estimation of the whole soil volume, we rely on the following assumption of *symmetry*: the portions of the soil volume that are not covered by the sensor grid behave as the monitored one; in other words, the profile of the uncovered portions can be obtained under rotation or translation of the moisture profile. Figure 5 shows the symmetries in the case of 2D and 3D grids. The symmetry assumption is verified for 2D grids when the drippers along a tree row are close enough to ensure homogeneous soil moisture. If the distance between the drippers is high and the soil moisture along the row is not constant, a 3D grid must be adopted to satisfy the symmetry assumption. Obviously, a 3D grid does not require any symmetry assumption if sensors are distributed all around the tree. However, given a fixed amount of sensors, not exploiting the assumption of symmetry results in a coarser grid.

We emphasize that all the soil moisture estimation approaches based on sensors implicitly make symmetry assumptions: if a single sensor (0D installation [9]) is used, the measured moisture is assumed as the reference value for the whole soil volume (i.e., it is assumed to be constant all over the soil

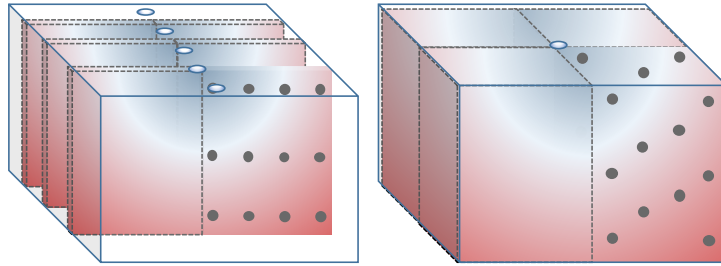


Figure 5: Symmetries in watered soil: dotted rectangles (left) or parallelepipeds (right) replicate the data collected by the sensor grids.

volume); if a column of sensors (1D installation [10, 11, 12, 13, 14]) is used the moisture is considered constant at a given depth. From this point of view, PLUTO's 2D and 3D grids progressively lighten such assumptions leading to a more accurate result. Obviously, some assumptions of symmetry remain, but they seem reasonable especially in the agricultural context: orchard rows are (or can be) built to be symmetrical. For example, the kiwi orchards involved in the Agro.Big.Data.Science project (i) were north-south oriented to have the same sun exposure on both the sides of the row; (ii) exploit a T-Bar training system; (iii) have a symmetric raised bed (i.e., with the same slope on both sides of the row).

More in general, to handle the variability within a single installation, it is always possible to design a grid layout that involves both sides of the row and, to handle variation within different parts of the field, two replicas can be installed.

#### 4.1. From Sensors to Moisture Profile

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

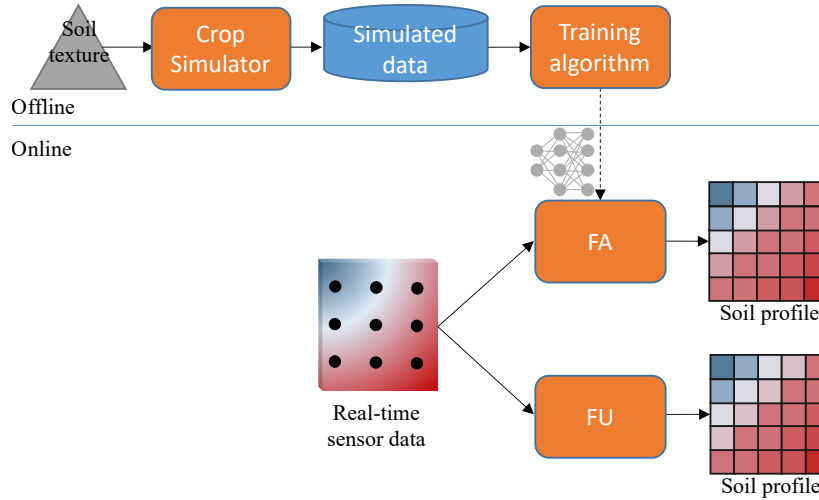


Figure 6: Feature-aware and feature-unaware processes for building a moisture profile.

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

The two alternative profiling functions are sketched in Figure 6. The feature unaware function does not require to be fitted to a specific field and only gets sensor data as input, while the feature aware function is trained offline to capture non-linear behaviors that result in a more accurate moisture profile. We emphasize that our feature unaware function also captures non-linear behaviors. Indeed using a grid of sensors, we are implicitly adopting a local regression approach [24, 25]. Although local regression approaches are bounded to linear behavior between sensor pairs, the composition of several linear strokes can approximate a non-linear trend. Conversely, our feature aware function models the non-linearities between sensor pairs too. The gap between the two approximations grows as the distance between the sensors increases.

As to the feature unaware profiling function, we rely on the well-known  $n$ -linear interpolation, where  $n$  is the profile grid dimensionality. 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 7). 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 7a), then (ii) we compute  $p.v$  (Figure 7b) 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. Below follows the formal definition.

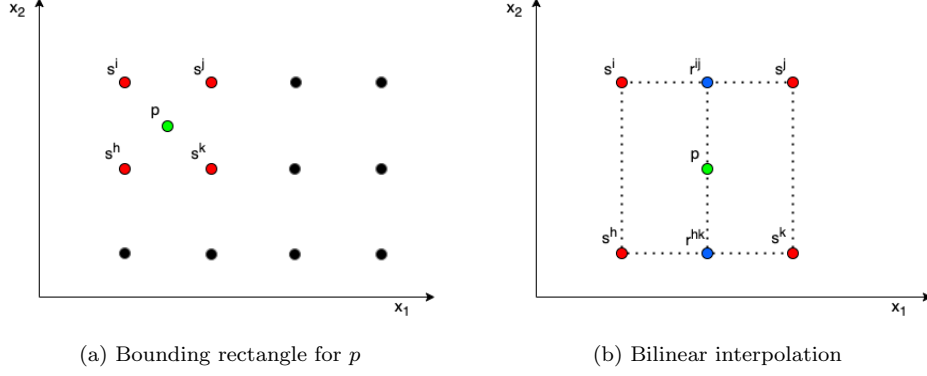


Figure 7: A 2D example of the feature unaware function.

$$\begin{aligned}
 r^{ij}.v &= \frac{s^j.x_1 - r^{ij}.x_1}{s^j.x_1 - s^i.x_1} s^i.v + \frac{r^{ij}.x_1 - s^i.x_1}{s^j.x_1 - s^i.x_1} s^j.v \\
 r^{hk}.v &= \frac{s^k.x_1 - r^{hk}.x_1}{s^k.x_1 - s^h.x_1} s^h.v + \frac{r^{hk}.x_1 - s^h.x_1}{s^k.x_1 - s^h.x_1} s^k.v \\
 p.v &= \frac{r^{ij}.x_2 - p.x_2}{r^{ij}.x_2 - r^{hk}.x_2} r^{hk}.v + \frac{p.x_2 - r^{hk}.x_2}{r^{ij}.x_2 - r^{hk}.x_2} r^{ij}.v
 \end{aligned}$$

The trilinear procedure is analogous: it just has three steps instead of two.

The feature aware function captures non-linear soil moisture behaviors by exploiting the knowledge about the hydrological fluxes. To this end, we rely on an Artificial Neural Network (ANN), a machine learning model, that returns the estimated values for each profile point after learning the soil behavior from the data generated by CRITERIA [20], a physically-based numerical model. The learning process is sketched in Figure 6. 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. CRITERIA reproduces hydrological fluxes in the soil volume implementing Richard's equations. The simulator models the following elements of the soil volume:

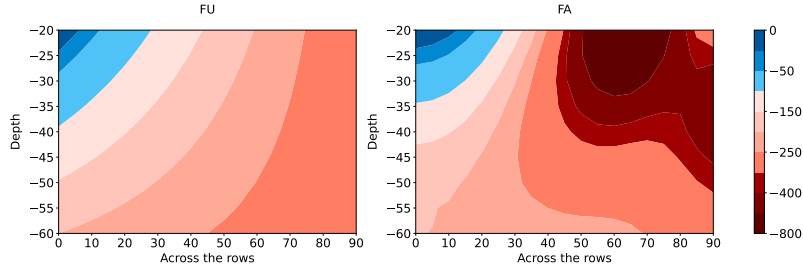


Figure 8: Moisture profile built by the two types of profiling functions, namely feature unaware (left) and feature aware (right).

(i) the soil texture; (ii) the shape of the tree roots; (iii) the watering system type and layout; and (iv) the sensor layout. Such data are provided as input by the farmer. Then, training data are generated running the simulator according to simulated weather conditions and watering sessions. Generating hourly data over a 4-month period proved to be sufficient to train the ANN. Refer to Section 8 for more details on the PLUTO’s ANN.

Figure 8 exemplifies the differences between the feature aware and feature unaware results. The figure shows two profiles obtained from a grid of four sensors positioned on the corners. The values for the sensors are the same for the two profiles (top-left=-10cbar, top-right=-300cbar, bottom-left=-200cbar, bottom-right=-300cbar). The profile on the left is obtained by applying the feature unaware bilinear profiling function (i.e., the 2-linear interpolation), while the one on the right by applying the feature aware ANN. The lower level of soil moisture in the upper right part of the feature aware profile results from the combination of two factors that are better captured by the feature aware function; this region (i) is proximal to the surface (thus it is more subject to atmospheric events) and (ii) belongs to the unwatered volume. We emphasize that this is a worst-case scenario: as already mentioned,

the higher the number of sensors, the lower the errors due to non-linear behaviors that are local to the intra-sensor regions. We analyze how accuracy changes for feature aware and feature unaware varying the number of sensors in Section 5.

#### 4.2. Visualization

Visualization allows easy exploitation of the moisture profile, enabling effective soil moisture monitoring. We propose two visual components selected according to the data types and the following visualization goals [26].

- *Historical soil moisture trend*: shows the variation of moisture over time. Independently of the profile dimensionality, a 2D stacked area chart is adopted. No spatial information is shown here.
- *Spatial soil moisture*: shows the spatial distribution of soil moisture at a specific time. A 2D heat map or a 3D scatter plot is adopted depending on the profile dimensionality.

The two components are linked through *interactive zooming* [27]. When the user selects a specific time on the historical soil moisture trend, the corresponding spatial soil moisture at the time is shown.

Since the goal is to check that an ideal moisture profile is maintained, we define five crop-specific soil moisture ranges, each associated with a different color. For our case of study on kiwi-trees [4], Dark blue (in  $[0, -30)$  cbar) and light blue (in  $[-30, -100)$  cbar) show heavily/slightly portions of over-watered soil. Salmon pink (in  $[-100, -300)$  cbar) represents the ideal case [28]. Finally, light red (in  $[-300, -1500)$  cbar) and dark red (in the range of  $[-1500, -\infty)$  cbar) show slightly/heavily portions of under-watered soil.

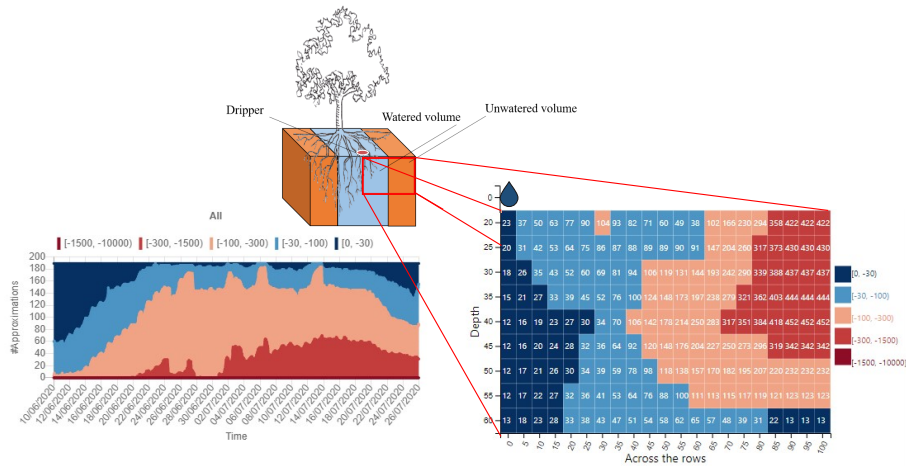


Figure 9: Visualization of a 2D profile: (left) the historical soil moisture and (right) a snapshot of the spatial soil moisture. The rightmost legend shows the soil moisture ranges. The drop at depth 0 represents the *Dripper* position and the rightmost legend shows the soil moisture ranges.

Figure 9 shows examples of historical soil moisture trend and spatial soil moisture charts for a 2D profile. The historical soil moisture chart shows how the soil gradually dries out during the dry season. In May the soil is mostly wet, while in June the first irrigations are necessary. Also, note how the single pipeline dripper fails to eliminate the red area due to the presence of an unwatered region in the soil volume. For the chosen zoom date, the spatial soil moisture chart shows the moisture profile in detail. Each cell corresponds to a profile element. The following areas can be highlighted: (i) the region under the dripper where the effects of irrigation are apparent; (ii) the superficial portion away from the dripper that is very dry as it is not irrigated; and (iii) the deep region which is less affected by atmospheric agents and in which constant soil moisture remains regardless of the effects

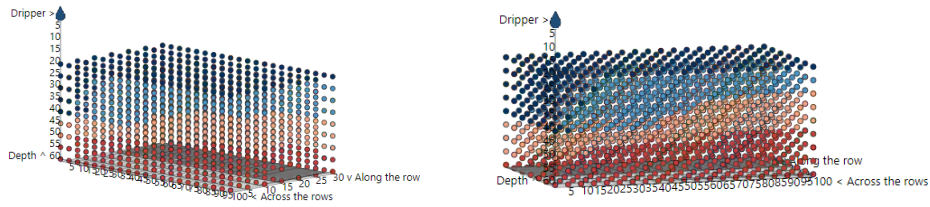


Figure 10: Visualizing rotations of a 3D profile.

of watering. The profile highlights how the soil moisture spreads laterally thanks to the combined effect of the roots and the permeability characteristics of the soil.

Figure 10 shows an example of a 3D spatial soil moisture chart. Each sphere corresponds to a profile element. The colors are the same as in the 2D visualization. Spaces between spheres allow the inner side of the 3D profile to be explored. Furthermore, to facilitate the analysis, the chart can be rotated along the three axes. The soil moisture gradient along the third dimension (i.e., along the row) is better highlighted by the third orientation of the chart and justifies the adoption of a 3D profile.

Starting from the spatial profile, we derive many meaningful visualizations. Figure 11-left shows the profile variance chart, which reports the soil moisture variance in a given period. The lighter areas are those where soil moisture varies the most. Similarly, Figure 11-right shows the profile average chart which reports the average soil moisture in a given period. The charts are designed to support the work of both agricultural technicians and farmers. For example, in the Agro.Big.Data.Science projects the main questions posed by such users were:

- *What is the watered volume?*: this region is typically characterized by

two phenomena: high moisture and high suction by the roots. Depending on the supplied quantity of water, the summation of the two phenomena determines different effects. In the case of over-watering, the watered volume remains always wet, with high values in the average chart and low values in the variance chart as the roots and the atmospheric phenomena are not able to absorb the water. Conversely, in the case of correct irrigation, the average chart shows medium values while the variance one shows high values mainly due to the suction made by the roots. This is the case for Figure 11.

- *Where is the root suction higher?:* a high root suction quickly reduces the moisture in the soil and results in a strong soil moisture variance.
- *How soil moisture dynamics impact on the watered volume?:* if, after increasing the quantity of water supplied to the soil, the watered volume does not increase significantly, then the soil disperses water in depth due to its characteristics.

In the context of the Agro.Big.Data.Science project agricultural technicians and some *data enthusiast farmers* [29] already had the competences to directly exploit the charts, while most of the farmers had been trained. We strongly believe that the progressive digitization in agriculture will drive all farmers to take advantage of these solutions, as it has already happened to operators in other industries.

A demo of the PLUTO visualization system is available at <https://big.csr.unibo.it/projects/pluto>. The monitoring system can be completed by the following information collected through a weather station and a water

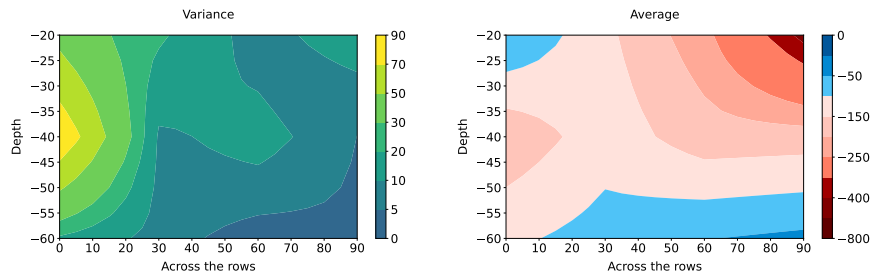


Figure 11: Soil moisture variance (left) and average (right) in cbar along a 1-week period; the orchard is watered through a single-pipeline dripper system.

meter:

- a single line chart showing the max/min/avg air temperature;
- a bar chart showing the mm of rain that fell each day;
- a bar chart showing the mm of water supplied, each day, through watering.

Although the role of data processing and visualization is often overlooked, such visualizations contribute to a comprehensive understanding of the evolution of the profile.

## 5. Results and Discussion

We tested PLUTO on the installations of the Agro.Big.Data.Science project [4]. Data have been collected in the irrigation season from May to September in 2020 and 2021. Installations are taken from two orchards: one in the plains and one in the hills around Faenza, in the province of Ravenna, Italy. The orchards were planted in 2010 as a self-rooting Hayward variety (*A. chinensis* var *deliciosa*), grafted in 2012 with Gold 3 (*A. chinensis* var



Location	# Grids	Grid Layout	# sensors	Watering system
Hilly	1	3D	15	Single pipeline
Hilly	2	2D	12	Single pipeline
Hilly	2	2D	12	Double pipeline
Plain	4	2D	12	Single pipeline

Table 2: Description of the experimental implants.

chinensis). Kiwifruit vines were spaced 2m along the row and 4.5m across the rows. Different irrigation systems and sensor grid layouts have been considered as shown in Table 2. Sensor data are sampled and collected every 15 minutes.

We emphasize that the number of sensors we use is much higher than necessary, as we will show below. Grids with 12-15 sensors were used to build the ground truth to carry out the tests. In each test only a subset of the available sensors was used to compute the profile, while the remaining ones were used as ground truth. In other words, we verify how well the profile approximates the soil moisture in positions where the real soil moisture is known to the tester but hidden to PLUTO. Profile accuracy is calculated as the Root Means Square Error (RMSE) between the ground truth sensor values and the values estimated by the profile. Obviously, the RMSE is computed only for those sensors that are not used to compute the profile.

### 5.1. Accuracy Evaluation

Figure 12 shows the system performance, for the 2D profile, varying the number of input sensors. Tests have been repeated for all the available in-

stallations. To better understand the advantage provided by PLUTO, we also highlighted 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. We refer to section Section 5.2 for an in-depth analysis of the layout of the profile sensors and the robustness of the system.

The 0D setting estimations are by far less accurate than the ones obtained by our profiling functions; indeed, three sensors are sufficient to halve its RMSE using the feature aware profiling function. The 1D setting achieves 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 [28].

RMSE for feature aware and feature unaware gradually decreases as the number of sensors increases. Note that the bilinear 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. This supremacy

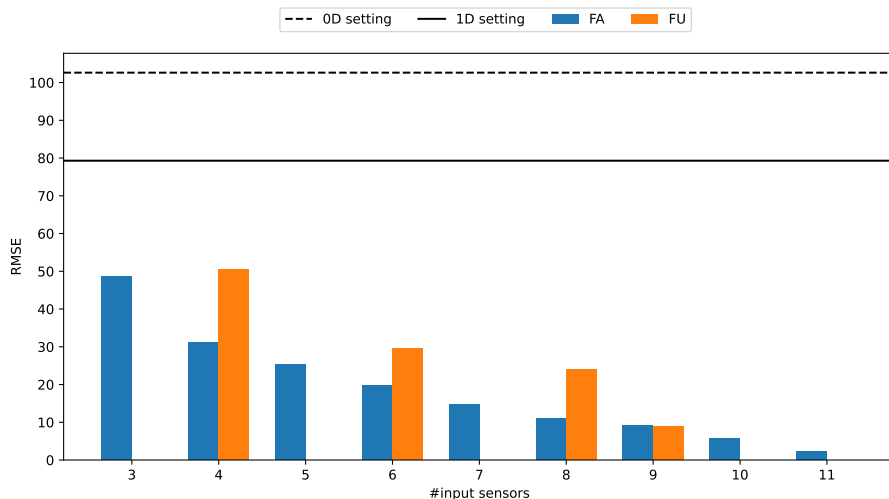


Figure 12: 2D profile performances as a function of the input sensors. Single sensor (0D) and column (1D) settings are reported for comparison.

is more evident when the number of sensors is limited and the intra-sensor distances are larger (see Figure 8).

3D profile performances are reported in Figure 13. This type of profile measures soil moisture variability along the row, while 2D profiles assume soil moisture to be constant along the row. The chart shows that the assumption is partially true since the RMSE for the 2D profile (i.e., the dotted line) is higher than those related to the 3D ones. The RMSE further increases for the 0D and 1D settings as they progressively make stronger assumptions about symmetry. As expected, the RMSE for the 3D profiles decreases as the number of sensors increases, and the feature aware profiling function overcomes the feature unaware one.

Overall, 2D and 3D profiles are much more accurate; they better capture variations in soil moisture than the baselines (that are traditionally used in

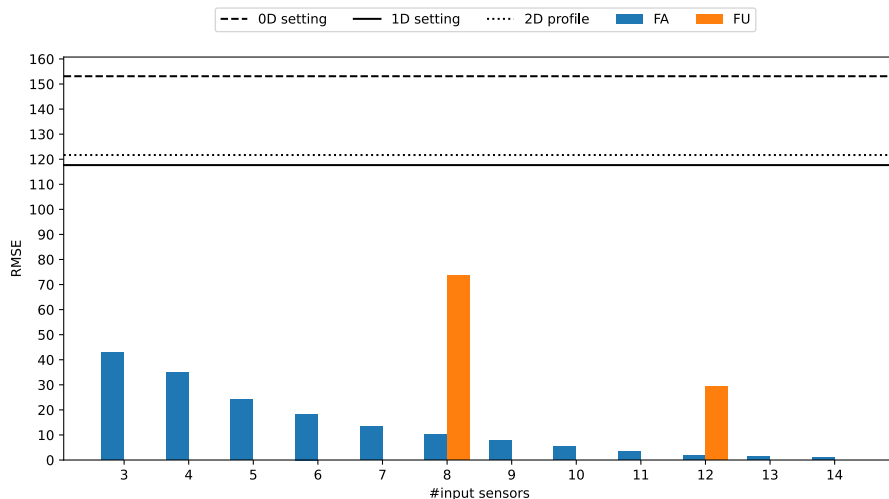


Figure 13: 3D profile performances as a function of the input sensors. Single sensor (0D), column (1D) settings, and (2D) profile are reported for comparison.

commercial settings). Figure 14 highlights the ability to identify different soil behaviors by showing the dynamics of soil moisture in fields with different watering systems. The orchard reported in Figure 14-left adopts a single-pipeline dripper system where the pipeline is placed along the row. The orchard in Figure 14-right adopts a double-pipeline dripper system where each pipeline is shifted by 20cm on the right/left of the tree row. As shown by the profile, the compound effect of the raised bed and the shift of the pipeline determines a heavily different soil moisture diffusion. This effect can only be pointed out using 2D or 3D profiles.

### 5.2. Sensor Layout Analysis

As shown in Figures 12 and 13, 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 feature unaware

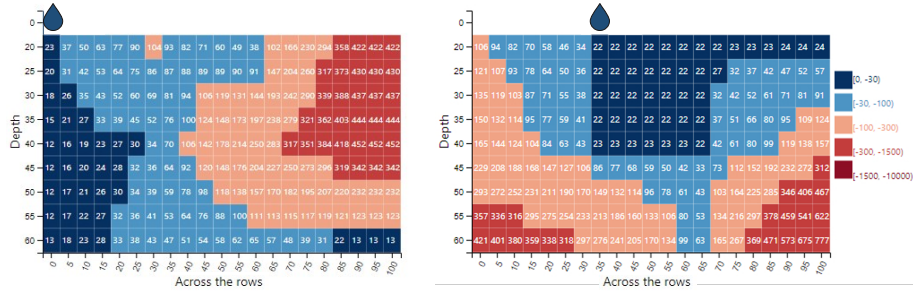


Figure 14: Profile comparison for orchards equipped with single pipeline dripper (left) and double pipeline dripper (right) watering systems. The drop at depth 0 represents the *Dripper* position and the rightmost legend shows the soil moisture ranges.

function implies some geometric constraints, in the ANN-based feature aware one all the layouts are feasible.

To compare different layout performances and the system stability with respect to different layouts, we considered the five profiles that achieve the best performances when  $m \in [3, 6]$  sensors are used. The analysis is carried out in an orchard watered through a single-pipeline dripper system monitored through a grid of  $n = 12$  sensors (i.e., 4 columns of 3 sensors). Figure 15 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. The sensors that turn out to be more relevant are:

- the sensor just under the dripper ( $0cm, -20cm$ ) since it is the most affected by the effects of the dripper;
- the sensor near the surface farthest from the dripper ( $90cm, -20cm$ ) since it records the state of the unwatered volume and is strongly in-

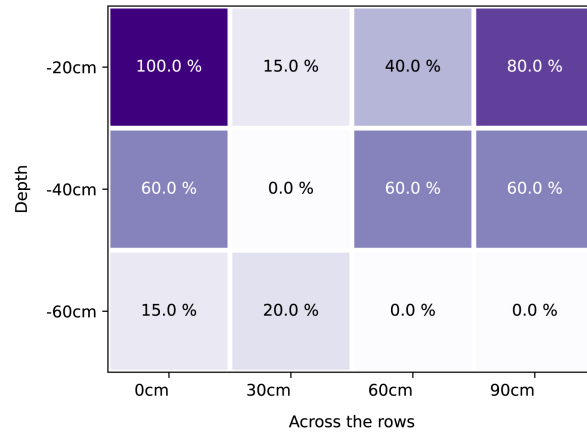


Figure 15: Frequency of times each sensor appeared in the best layouts.

fluenced by atmospheric phenomena (e.g., sun, rain, and wind);

- one sensor at mid-depth (\*,  $-40cm$ ) since it captures the soil behavior when not directly affected by watering and atmospheric phenomena.

Besides proving the stability and robustness of the system, this analysis provides the rules for defining the sensor layout.

### 5.3. Application of PLUTO in Improving Water Management and Fruit Quality

We briefly report the results achieved in the Agro.Big.Data.Science project in terms of water saving and fruit quality. Although an in-depth discussion of such results is out of the scope of this paper, we believe that it is important to demonstrate the ultimate impact of PLUTO; for more details see [30].

We have two irrigation setups during the 2021 campaign (i.e., from May to October 2021) within the same orchard: (i) managed row (irrigation is *automatically* controlled by PLUTO) and (ii) control row (irrigation is *manually* controlled by the farmer). As to (i), given a 2D installation of 12 sensor, the moisture profile was exploited in the water-saving strategy: irrigation was activated when the soil water content got down below the field capacity ( $-0.003\text{MPa}$ ) and returned the same amount of water lost to evapotranspiration the day before (as measured by a PAN evaporimeter [30]).

As to water saving, the managed row saved 44% of water during the whole campaign. We achieved the maximum saving in June and September when, for the farmer, it is more difficult to accurately estimate the actual soil moisture level and the water requirement.

As to fruit quality, the productivity of vines was unaffected by the irrigation management and ranged from 32 to 39 kg/vine (35-44 t/ha). Fruits from the control row appeared greener with a hue angle of 105 vs a hue angle of 102 for fruits from the managed row. Such fruits also showed a higher soluble solid concentration at harvest: 15.3 brix vs 12.7 brix for the control row. The gap has been maintained after 2 months of storage (and 1 day of shelf life); in particular, the soluble solid concentration was 17.4 brix for the managed rows vs 16.1 brix for the control row.

## 6. Conclusions and Future Work

We presented PLUTO, an original approach to compute 2D/3D moisture profiles with granularity in the order of magnitude of a few square/cubic centimeters. PLUTO relies on a grid of soil moisture sensors and it largely

outperforms previous approaches based on a single or a column of sensors. To create a cost-effective operative solution, we have shown that three sensors, properly placed in the soil, are sufficient to effectively obtain the profiled soil moisture.

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 forecast module fed with the weather forecast.

## **7. Acknowledgments**

Thanks to Prof. Moreno Toselli, Elena Baldi and Maurizio Quartieri for their support and for sharing their knowledge on kiwi cultivation.

## **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 (8) (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 (1) (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 (2) (2008) 587–600. doi:10.2136/vzj2007.0077.
- [4] The Agro.Big.Data.Science project, <http://agrobigdatascience.it/>, last accessed: 2021-10-18.
- [5] M. Bordoni, M. Bittelli, R. Valentino, S. Chersich, M. Persichillo, C. Meisina, Soil water content estimated by support vector machine for the assessment of shallow landslides triggering: the role of antecedent meteorological conditions, *Environmental Modeling and Assessment* 23 (4) (2018) 333–352. doi:10.1007/s10666-017-9586-y.
- [6] A. Tu, S. Xie, M. Mo, Y. Song, Y. Li, Water budget components estimation for a mature citrus orchard of southern china based on hydrus-1d model, *Agricultural Water Management* 243. doi:10.1016/j.agwat.2020.106426.
- [7] G. Hamilton, G. Akbar, S. Raine, A. McHugh, Deep blade loosening and two-dimensional infiltration theory make furrow irrigation predictable, simpler and more efficient, *Agricultural Water Management* 239 (2020) 106241.
- [8] 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. doi:10.1016/j.rse.2021.112434.

- [9] 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.
- [10] F. Karandish, J. Šimůnek, A comparison of numerical and machine-learning modeling of soil water content with limited input data, *Journal of Hydrology* 543 (2016) 892–909. doi:10.1016/j.jhydro.2016.11.007.
- [11] 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 (3) (2018) 421–444. doi:10.1007/s11119-017-9527-4.
- [12] 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 (5) (2020) 1327–1348. doi:10.13031/TRANS.13765.
- [13] 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. doi:10.1016/j.agwat.2020.106542.
- [14] 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 (5) (2015) 2382–2401. doi:10.3390/w7052382.

- [15] G. Egea, A. Diaz-Espejo, J. Fernández, Soil moisture dynamics in a hedgerow olive orchard under well-watered and deficit irrigation regimes: Assessment, prediction and scenario analysis, *Agricultural Water Management* 164 (2016) 197–211. doi:10.1016/j.agwat.2015.10.034.
- [16] 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 (1) (2016) 139–148. doi:10.7451/CBE.2015.57.1.39.
- [17] 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 (5). doi:10.3390/w13050600.
- [18] 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 (6). doi:10.3390/w13060791.
- [19] G. Vitali, M. Francia, M. Golfarelli, M. Canavari, Crop management with the iot: An interdisciplinary survey, *Agronomy* 11 (1) (2021) 181.
- [20] 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.
- [21] M. Chen, G. R. Willgoose, P. M. Saco, *Spatial prediction of temporal*

- soil moisture dynamics using hydrus-1d, *Hydrological processes* 28 (2) (2014) 171–185.
- [22] E. V. Shein, A. Y. Mady, L. I. Il'in, Validation of hydrus-1d for predicting of soil moisture content with hysteresis effect, *Biogeosystem Technique* (2019) 59–64.
- [23] A. Hinnell, N. Lazarovitch, A. Furman, M. Poulton, A. Warrick, Neurodrip: Estimation of subsurface wetting patterns for drip irrigation using neural networks, *Irrigation Science* 28 (6) (2010) 535–544. doi:10.1007/s00271-010-0214-8.
- [24] C. Loader, *Local regression and likelihood*, Springer Science & Business Media, 2006.
- [25] W. Cleveland, S. Devlin, Locally weighted regression: An approach to regression analysis by local fitting. *journal of the american statistical association*, vol., 83, 596.
- [26] M. Golfarelli, S. Rizzi, A model-driven approach to automate data visualization in big data analytics, *Information Visualization* 19 (1) (2020) 24–47.
- [27] D. A. Keim, *Information visualization and visual data mining*, *IEEE transactions on Visualization and Computer Graphics* 8 (1) (2002) 1–8.
- [28] S. Miller, G. Smith, H. Bolding, A. Johansson, Effects of water stress on fruit quality attributes of kiwifruit, *Annals of botany* 81 (1) (1998) 73–81.

- [29] K. Morton, M. Balazinska, D. Grossman, J. Mackinlay, Support the data enthusiast: Challenges for next-generation data-analysis systems, *Proceedings of the VLDB Endowment* 7 (6) (2014) 453–456.
- [30] M. Quartieri, M. Toselli, E. Baldi, G. Polidori, M. Germani, M. Noferini, E. Xylogiannis, Effect of the method and volume of irrigation on yield and fruit quality of yellow fleshed kiwifruit in northern italy, in: *X International Symposium on Kiwifruit 1332*, 2021, pp. 211–218.
- [31] 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 (1) (2015) 014008.

## 8. Appendix: the PLUTO’s Artificial Neural Network

ANNs take inspiration from biological neural networks: they encompass layers of neurons, linked to each other with connections. Connections behave as synapses in a biological brain, with certain intensity (called weight). The input is fed to the first layer (called input layer), processed throughout the remaining layers (called hidden layers), and finally provided in the last layer (called output layer) as an outcome. ANNs are trained by feeding inputs and the expected outputs a given number of times (i.e., epochs) in groups of a certain size (i.e., batch size). Such a process allows the network to tune the weights of the connections that link the neurons together, and find the optimal solution.

ANNs have been chosen for the feature aware profiling function due to the following properties:

Table 3: The ANN hyper-parameter configuration retrieved by HyperOpt.

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

- Low resource consumer: once trained offline, the ANNs are fast and require limited computational resources.
- Powerful & Flexible: can provide accurate profiles (i.e., they properly capture non-linearities) even though the grid layout includes a limited number of sensors.
- Robust: ANNs are well-known to work well even in case of dirty or missing data. This may happen when a sensor occasionally fails to provide the correct data.

Table 3 reports the hyper-parameters<sup>2</sup> for the PLUTO ANN. We adopted a feed forward ANN (i.e., neurons in the layers are connected without creating any cycle) with one, fully connected, hidden layer. The input layer

---

<sup>2</sup>Hyper-parameters are the variables that determine the network structure (e.g., number of hidden layers, number of neuron per layer), how the network is trained (e.g., number of epochs, batch size), and hence the final performance.

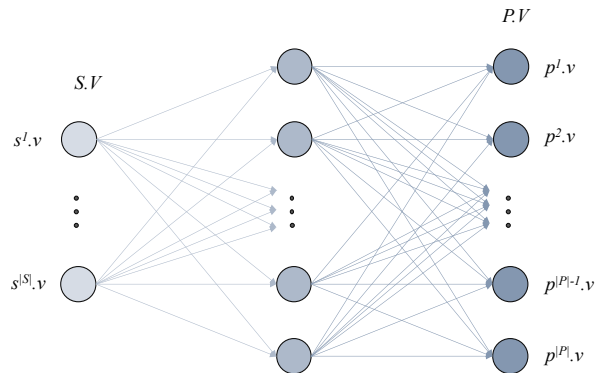


Figure 16: The layout of the ANN profiling function.

has one neuron for each sensor, the hidden layer has a hundred of neurons activated through the hyperbolic tangent activation function (e.g., values of the neurons are mapped according to the homonym mathematical function), and the output layer has as many neurons as the number of points in  $P$  (see Figure 16). As to the learning hyper-parameters, the process involves 50 epochs, a batch size of 30 samples, and the learning rate (i.e., the step size towards the final solution) is reduced by a factor of 2 when no improvement is seen for a *patient* number of 10 epochs. Besides, a Z-score normalization (i.e., transforming data into a Gaussian distribution with mean 0 and standard deviation 1) is applied to the input values before feeding the network. This is due to the following reasons.

- To give all the input values the same importance. Even though the soil moisture varies everywhere in the same range, we noticed sensible variations depending on the location of the sensor (e.g., the area right under the dripper and the opposite corner).
- To address ANN-specific training issues. Normalizing the inputs re-

duces the chances of getting stuck in local optima (e.g., balance the rate at which the ANN learns) and makes training faster (e.g., avoid the saturation of non-linear activation function).

All the hyper parameters (i.e., ANN structure and learning rates) have been set through a hyper-parameter tuning process implemented with HyperOpt [31]. HyperOpt exploits state-of-the-art optimization techniques to heuristically explore the huge search space of hyper-parameters. The adoption of a training set including several different weather conditions and watering sessions avoids overfitting (i.e., a learning process that returns a network that works well just for the specific cases it was trained for).