AgriEngineering

MDPI

Article

Letting the Crop Speak: A Proof-of-Concept Sensor-Grounded Large Language Model Advisor for Potato Irrigation on a Mediterranean Vertisol

Ilyass ElAlaoui *, Asmaa Monafii, Doha Boumalik, Younes Aourir, Assia Tahiri, Ghoufrane Gouddar and Oussama Hassani

Natural Resources and Sustainable Development Laboratory, Department of Biology, Faculty of Sciences, Ibn Tofail University, Kenitra 14000, Morocco; asmaa.monafii@uit.ac.ma (A.M.); doha.boumalik@uit.ac.ma (D.B.); younes.aourir@uit.ac.ma (Y.A.); assia.tahiri@uit.ac.ma (A.T.); ghoufrane.gouddar@uit.ac.ma (G.G.); oussama.hassani@uit.ac.ma (O.H.)

* Correspondence: ilyass.elalaoui@uit.ac.ma

Abstract

Low-cost soil sensors are now widespread, but turning a moisture reading into an irrigation decision still requires expertise that many growers lack. This proof-of-concept study describes a conversational advisor in which a general-purpose large language model (GPT-5) receives live readings from a low-cost capacitive probe, together with a short crop and soil profile, and answers the operator in the first person, as the plant. The advisor was used throughout a 14-week greenhouse potato (*Solanum tuberosum* L.) cycle in a clay-rich Moroccan Vertisol ("Tirs"). Six pots were irrigated only on its recommendation and six by an experienced technician working without sensor data. The advisor-guided block received less water, and its plants grew taller and set more tubers of larger maximum diameter. Because the treatments were neither replicated nor randomised, irrigation was logged per block and tuber mass was not recorded, these observations are descriptive and cannot be attributed to the language model rather than to the sensor, bench position or irrigation schedule. The recorded dialogue showed the advisor reasoning faithfully from its inputs, including a mislabelled moisture unit. The study demonstrates technical feasibility, not the agronomic effectiveness of the recommendations, and defines the controls a replicated trial requires.

Keywords: large language models; sensor grounding; conversational decision support; precision irrigation; *Solanum tuberosum*; Vertisol; human–AI interaction; proof of concept

Check for updates

Academic Editor: Ernane José Xavier Costa

Received: 30 September 2026
Revised: 25 October 2026
Accepted: 28 October 2026
Published: 3 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Agriculture accounts for roughly 70% of global freshwater withdrawals, and in many river basins the water that remains is already committed several times over [1]. Morocco sits near the sharp end of this problem. A succession of dry years has drawn down reservoirs and aquifers, and climate projections point to longer and more frequent droughts across the country's main agricultural regions [2]. Potato is one of the crops most exposed to that trend. It is a staple of Moroccan vegetable production [3], yet its shallow, relatively sparse root system makes it one of the least drought-tolerant field crops, and even brief water deficits during tuber initiation and bulking reduce both tuber number and size [4]. Potato is just as sensitive to excess water, so the range of soil water tension within which it

performs well is narrow [5]. Irrigating potato well, in other words, is less a matter of giving more water than of giving it at the right moment.

The hardware to monitor that moment is no longer the bottleneck. Capacitive soil moisture probes and Wi-Fi microcontrollers now cost a few dollars, and a large body of work has shown how they can be assembled into Internet of Things (IoT) networks for field monitoring [6–8]. What these systems deliver, however, is a stream of numbers. Knowing that the volumetric water content is 18% is useful only to someone who also knows what 18% means for a given crop, at a given stage, in a given soil. That knowledge is unevenly distributed, and the gap between the availability of digital tools and their use on farms is well documented [9,10]. Low-cost sensors also need soil-specific calibration before their readings can be taken at face value [11], which adds another layer of interpretation the grower has to carry.

The established response has been to automate the decision. Controllers that open a valve when a soil water reading crosses a set point have been available commercially for decades and can save substantial water [12,13], and machine learning (ML) models trained on sensor and weather histories now forecast irrigation need [14–16]. Some of these approaches are fully transparent; a threshold rule is easy to state and to audit. What they share is that their decision logic is fixed in advance. A threshold has to be chosen for each crop, stage and soil, and an ML model usually has to be retrained before it transfers to new conditions. Neither can be asked why it recommends what it does, in terms of the crop's situation, or answer a question its designer did not anticipate. For growers who want to understand a recommendation rather than simply follow it, that is a real limitation.

Large language models (LLMs) offer a different route. GPT-4 already answers agronomy certification exams at or above the level of human candidates [17], and LLMs are being tested as a way to extend the reach of agricultural advisory services [18,19]. Most of this work, however, treats the model as a knowledgeable conversational partner that knows nothing about the actual field. Its answers are only as specific as the question, and field tests have found them too generic to act on [18]. The obvious next step is to ground the model in live measurements, and the first systems to do so have appeared. They differ from the present one in ways that matter. IoT-LLM restructures sensor streams so that LLMs can reason over them, but it was evaluated on benchmark reasoning tasks rather than on a crop [20]. AgriBot coupled an ESP32 node to an LLM bound by explicit control rules and was tested for seven days, without measuring how the crop responded [21]. LLM agents have also scheduled a canal network, but in a hydraulic simulator [22]. To our knowledge, no study has yet run a sensor-grounded LLM advisor through a full crop cycle on living plants, alongside a person managing the same crop.

This study is a proof of concept of such an advisor. We connected a low-cost root-zone probe to GPT-5 through a context layer that combines the live reading with a description of the crop, its growth stage, its soil and its container, and asked the model to answer as the plant itself. We call the system a sensor-grounded conversational advisor, and its first-person voice a crop persona. We used it to manage six potato pots over a full cycle in a Moroccan Vertisol, a soil whose shrink–swell behaviour and slow internal drainage make visual judgement unreliable, while an experienced technician managed six other pots without sensor data. The advisor-guided pots therefore had both the sensor and the language model, and the technician's pots had neither. The comparison describes the net outcome of the "sensor plus LLM" package against unaided experience. It cannot show how much the language model adds over the same sensor read by a fixed threshold rule, or by a technician who is given the readings. We asked three questions: (i) Can such an advisor be built cheaply and operated over a full potato cycle? (ii) How do water applied, plant growth and tuber set in the advisor-guided block compare, descriptively, with those in the

technician-managed block? (iii) What does the recorded dialogue reveal about how such a system can fail? The contributions are a low-cost architecture documented down to its full system prompt (Supplementary File S1), a full-season case record of its use, and an analysis of the failure modes the dialogue exposed, which leads to a set of design safeguards and to the controls that a replicated trial will need.

2. Literature Review

2.1. From Sensing to Scheduling

Precision irrigation research has developed along two lines. The older and more widely adopted line schedules irrigation from soil water thresholds. Soil water tension or water content is measured, and irrigation is triggered when the reading crosses a crop-specific limit. The limit may be derived from the allowable depletion fraction tabulated in FAO-56 [23], or calibrated in the field against plant water status [5,24]. Automated versions close the loop. Tensiometer and granular-matrix controllers have run drip irrigation of tomato without operator intervention [25], commercial soil moisture sensor controllers have been tested for precision under field conditions [12], and reviews of automatic irrigation control describe a range that extends from simple on–off rules to model-predictive and fuzzy controllers [13]. These systems are the natural benchmark for any sensor-based advisor. If a fixed threshold applied to the same probe reading led to the same decisions, a language model would add nothing but an interface.

The second line is the data-driven stack that grew around the Internet of Things. It began with instrumentation, showing that distributed soil and climate sensors could be deployed and networked reliably at low cost [6,7]. It continued with platforms that aggregate the data in the cloud and present them to users or pass them to control logic. The SWAMP project, for example, demonstrated IoT-based precision irrigation across several European and Brazilian pilots [26], and reviews have catalogued the benefits and obstacles of agricultural data analytics more broadly [8]. Prediction came next. Goap et al. [14] combined soil moisture sensing with weather forecasts in an ML model that anticipates irrigation need, Vij et al. [15] automated farm irrigation with a similar combination, and Benos et al. [16] reviewed how widely such models are now used. On both lines, the output is either a dashboard that a person must interpret or a control action taken by fixed logic. The step of explaining a reading in terms of the crop's situation, in a form a non-specialist can question, has received much less attention.

2.2. Digital Twins and Conversational Advisors

Digital twins, virtual counterparts that are continuously updated from their physical objects, have attracted growing interest in agriculture as a way to monitor, predict and control farm systems [27,28]. The term has a fairly precise meaning. A digital twin exchanges data automatically in both directions with its physical object and holds a model that mirrors, and can simulate, the object's state. A system in which data flow automatically only from the object to the model is better described as a digital shadow [29]. By that definition, the system described here is not a digital twin. Readings pass automatically from the probe to the language model, but the model keeps no state between conversations and simulates no crop physiology, and every action on the plant passes through a person. We therefore call it a sensor-grounded conversational advisor. What is new about it is how it communicates rather than what it simulates: the latest measurement is rendered as a first-person account from the plant, which the operator can question in ordinary language.

2.3. Large Language Models in Agriculture

The first wave of agricultural LLM studies tested what the models know. Silva et al. [17] reported that GPT-4 passed agronomist certification exams from Brazil, India and the United States, and that retrieval-augmented generation improved its accuracy further. Zhao et al. [30] showed that ChatGPT could classify agricultural texts across languages without task-specific training, and Zhu et al. [31] reviewed the growing range of language and vision models applied to crop management, pest diagnosis and advisory work. Studies of practical use have been more cautious. When GPT-4 was asked to advise Nigerian cassava farmers, its answers were fluent but often too generic, or wrong on timing, to be used as they stood, which led Tzachor et al. [18] to argue for keeping human experts in the loop. Mourtzinis et al. [19] reached a similar conclusion for soybean management recommendations: trust and field specificity mattered as much as accuracy.

A second wave has tried to anchor the models in real data and give them tools. Retrieval-augmented generation grounds answers in documents fetched at query time [32]. IoT-LLM restructures raw sensor streams into a form that LLMs can reason over, which improves their performance on physical-world reasoning benchmarks [20]. Agent frameworks go further and let the model plan and call external tools [33]. In irrigation specifically, Saranya et al. [21] coupled an ESP32 node to an LLM constrained by control rules and reported water savings of about 37.5% relative to manual scheduling over a seven-day test, but did not measure crop growth or yield. Chen et al. [22] used LLM agents to schedule a multi-level canal network in a hydraulic simulator and found that the physical constraints had to be written into the prompt before the models handled the numbers reliably. Table 1 places the present study among these efforts. None of them followed a crop through a season, and none compared the system with a person managing the same crop. Those are the gaps a proof of concept can begin to fill, although a single unreplicated trial can only describe them, not measure them.

Table 1. Position of this study relative to recent work on language models in agricultural decision support.

Study	Live Sensor Input	Role of the Language Model	Evaluation	Crop Outcome Over a Season
Silva et al. [17]	No	Answers agronomy exam questions (GPT-4, with retrieval)	Exam accuracy against human candidates	No
Tzachor et al. [18]	No	Drafts advisory messages for cassava farmers	Review by agronomists and extension staff	No
Mourtzinis et al. [19]	No	Soybean management recommendations from the literature	Expert rating	No
An et al. [20]	Yes	Reasons over preprocessed IoT sensor data	Benchmark reasoning tasks	No
Saranya et al. [21]	Yes (ESP32 node)	Rule-constrained irrigation advice	7-day test against manual scheduling	No
Chen et al. [22]	Simulated (hydraulic model)	Agent scheduling canal gates	Stability and goal achievement across models and prompts	No
This study	Yes (ESP32 node)	First-person crop persona advising irrigation	14-week unreplicated case study alongside an experienced technician	Descriptive only (water, height, tuber number and maximum diameter)

2.4. Risks of Conversational Advisors

Two bodies of work outside agriculture bear directly on a system of this kind. The first concerns hallucination, the tendency of generative models to produce fluent statements that are unsupported by their input [34]. In an irrigation advisor, a hallucinated volume or an ignored reading translates directly into water wasted or a crop stressed. The second concerns how people respond to automated advice. Operators who come to rely on an automated aid tend to check it less, a pattern described as automation complacency and automation bias [35]. Anthropomorphism, giving a system human-like features such as a first-person voice, is known to increase engagement and perceived competence [36,37]. Giving the crop a voice may therefore make the advice more persuasive, whether or not it is right. We return to this tension in Section 5.3.

2.5. Why a Vertisol and Potato Make a Demanding Test

Vertisols are dominated by swelling clays. They crack deeply as they dry and seal as they wet, so their hydraulic behaviour changes with water content and differs sharply between the matrix and the crack network [38]. The Tirs soils of the Gharb plain are a regional example. Their structure and water behaviour are highly sensitive to management [39], and a dry, cracked surface can conceal substantial water at depth. The least limiting water range, the band of water content within which root growth is limited neither by water stress nor by poor aeration or high soil strength, is typically narrow in such soils [40]. Excess water quickly produces hypoxia in the root zone [41]. Potato adds its own narrow tolerance. FAO-56 assigns it one of the lowest allowable depletion fractions among field crops [23]. The response of potato to reduced irrigation depends on the soil, on the growth stage at which the deficit occurs and on how water is applied. Some studies report that moderate reductions can be tolerated with little loss of yield, while deficits during tuber initiation and bulking are consistently costly in yield and quality [42–44]. A system that irrigates potato on a Vertisol therefore has little margin for error in either direction, which makes the combination a demanding setting in which to try out an advisor.

3. Materials and Methods

3.1. Site and Growing Conditions

The trial was conducted in the research greenhouse of the Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco, under a Mediterranean climate with marked seasonal variation in temperature and humidity. The greenhouse relied on natural light and passive ventilation, so plants experienced the ambient temperature and radiation regime of the region rather than a controlled one. Air temperature and humidity inside the greenhouse were not logged.

3.2. Plant Material and Soil

Certified seed tubers of potato (*Solanum tuberosum* L.), treated for phytosanitary compliance, were planted one per pot in 5 L pots. About three to four weeks after planting, when the plants had outgrown these containers, all twelve plants were transplanted at the same time into 25 L pots filled with the same substrate. The substrate was collected in the Oulad Ben Hammadi area of the Gharb plain (34.227073° N, 5.861078° W). It is a Chromic Vertisol, known locally as Tirs. Soils of this type are dark, rich in swelling clay (typically 40–50%) and markedly shrink–swell; they are fertile and retain water well, but their low hydraulic conductivity makes them prone to waterlogging when irrigation is poorly timed [38,39]. No laboratory analysis of the substrate used in this study (particle size distribution, bulk

density, pH or water retention curve) was carried out, so these general properties are taken from the literature on Tirs soils rather than measured.

3.3. Sensing Node

A dedicated sensing node was built for the study (Figure 1). It is based on an ESP32 microcontroller (Espressif Systems, Shanghai, China), chosen for its dual-core processor and integrated Wi-Fi, and is connected to an industrial two-in-1 capacitive probe (model JXBS-3001-TR, Weihai Jingxun Changtong Electronic Technology Co., Ltd., Weihai, China) that measures volumetric water content (VWC) and soil temperature simultaneously. Capacitive rather than resistive sensing was used to avoid electrode corrosion in the clay. The electronics are housed in a weather-resistant IP65 enclosure designed and 3D-printed in-house and powered by a generic rechargeable lithium-polymer battery, so the node can be carried from pot to pot. Readings are pushed over Wi-Fi to the cloud backend described below.

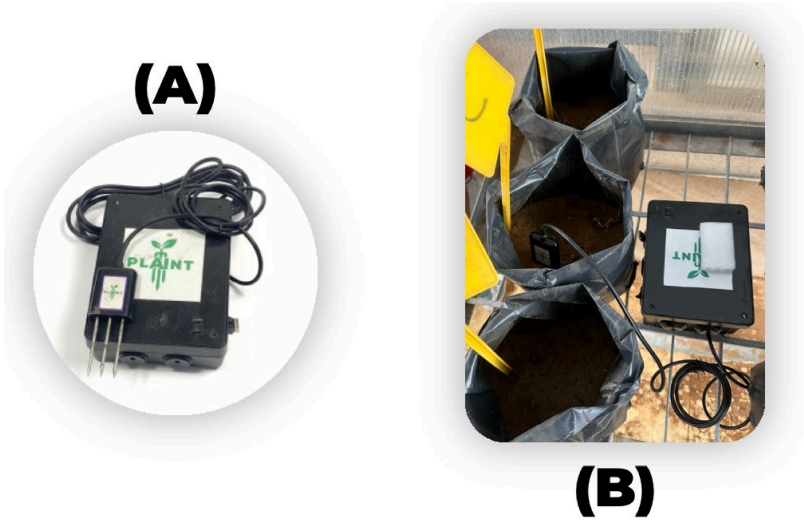


Figure 1. The sensing node: (**A**) encased device with the capacitive two-in-1 probe; (**B**) the probe inserted in a potato pot during a measurement.

The probe was used with the manufacturer's factory calibration; no soil-specific calibration was performed. The output of capacitance sensors depends on soil texture, electrical conductivity and temperature [45], and factory equations are least reliable in clay-rich soils such as this one [11]. The readings reported here should therefore be read as a relative index of wetness in this soil rather than as accurate absolute values of volumetric water content. In particular, the accuracy of the 8% reading discussed in Section 4.4 is unknown.

3.4. Context-Grounded Language Model

The software layer is a web application backed by Google Firebase (Google LLC, Mountain View, CA, USA; Figure 2). Before monitoring begins, the operator registers each plant by entering a short profile: a nickname, the species, the current phenological stage (for example, tuber initiation or bulking), the soil type and an optional free-text field for additional context (Figure 3). The operator updates the growth stage as the crop advances. When a reading arrives, the backend stores it alongside the profile.

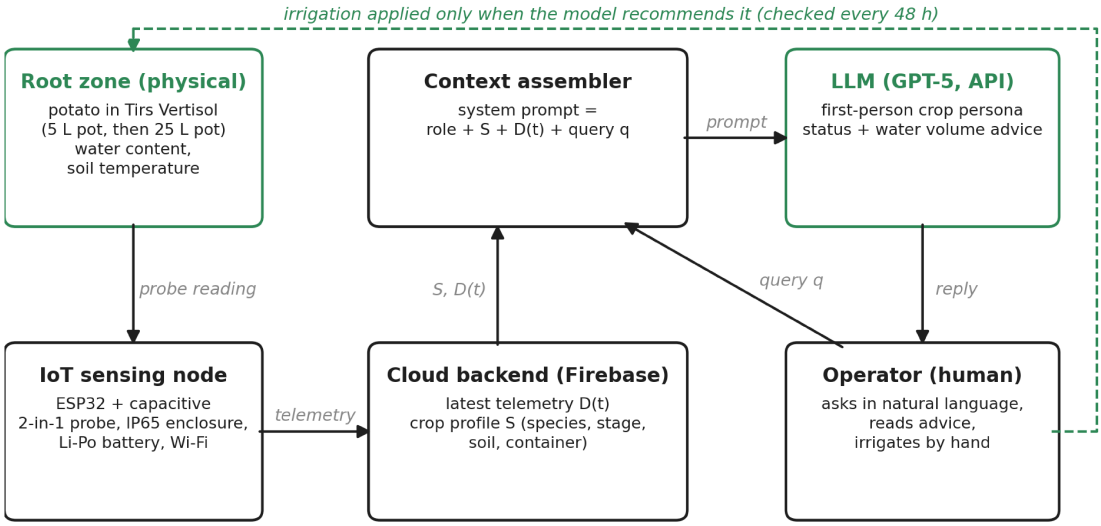


Figure 2. Architecture of the conversational irrigation advisor. S, crop profile; $D(\mathrm{t})$, latest sensor reading; q, operator query. Green outlines mark the crop and its first-person language-model persona; black outlines mark the hardware and software components. The dashed line closes the loop through the human operator.

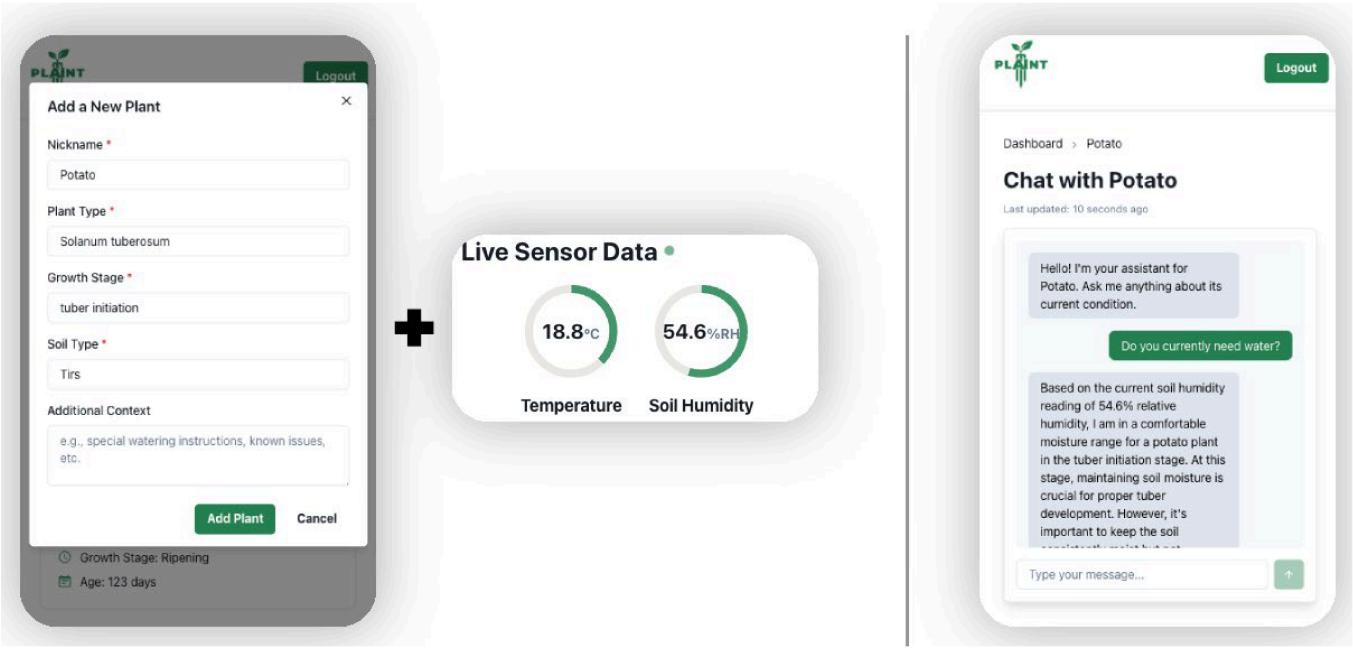


Figure 3. Application interface. (**Left**) Plant registration combined with live sensor readings. (**Right**) Conversational view in which the plant reports its status and water needs.

When the operator opens a conversation, a context assembler builds the prompt sent to the language model. The prompt, P, at time t can be written as

$$P(t) = \langle R, S, D(t), q \rangle \quad (1)$$

where R is a fixed role instruction, S is the crop profile, $D(t)$ is the latest sensor reading and q is the operator's question. In the software these four parts are joined into a single system message, reproduced in full in Supplementary File S1. The role instruction asks the model to analyse the readings in light of the crop's species, stage, soil and container, to answer the operator's question directly, to suggest a specific action when conditions are not ideal, and to reply in the first person as the plant. Replies therefore take the form "my soil is..." rather

than "the soil is…". The container was described in a fixed line of the message. While the plants were in 5 L pots it read "Medium-sized pot (approx. 5 L)"; after transplanting it was changed to describe a large pot of approximately 25 L. The water reading was inserted with the label "% RH", a unit of air humidity, although the probe reports volumetric water content. We return to this labelling error in Section 4.4.

The prompt contained no irrigation thresholds, soil water limits, depletion fractions or rules for computing volumes. The decision to irrigate and the volume suggested were produced by the model from the reading, the profile and its general training, not from rules that we supplied. The replies were generated by GPT-5 (OpenAI, San Francisco, CA, USA) [46] through the Chat Completions API, using the model identifier gpt-5. No dated snapshot was fixed in the request, so each call used the snapshot served for that identifier at the time. The completion length was limited to 200 tokens. Temperature and the other sampling settings were not specified and therefore followed the API defaults. Each request contained the system message, followed by the earlier turns of the current conversation and the operator's latest question. The chat opens with a fixed greeting written by the interface; it is not generated by the model, but it is passed back with later turns as an earlier assistant message. Model responses were displayed to the operator but not logged. Because neither the snapshot nor the sampling settings were fixed, the same input could have produced different recommendations on different calls [47], and individual recommendations made during the trial cannot be reconstructed. The absence of a complete interaction record limits more than reproducibility. It also prevents any assessment of decision consistency over the season, for example whether similar readings at similar growth stages led to similar decisions and volumes, so the consistency of the advisor's recommendations throughout the trial could not be evaluated. The model had no control over hardware. Every recommendation passed through the operator.

3.5. Experimental Design

The twelve plants were divided into two groups of six, one per treatment, placed on opposite halves of a single greenhouse bench (Figure 4). Pot positions were neither randomised nor rotated during the season. Each treatment therefore forms a single block, and bench position cannot be separated from treatment. The six pots within a block are subsamples of that block rather than independent replicates of the treatment. Light, temperature and ventilation commonly vary systematically along greenhouse benches [48], and plant height and tuber set respond to such gradients, so this layout is the main constraint on the interpretation of the results.

In the expert block, the greenhouse's agricultural technician decided when and how much to irrigate from visual inspection of the plants and soil and from his usual schedule. He had no access to sensor data and could irrigate at any time. The LLM-guided block was managed only through the advisor, following the protocol in Section 3.6. Apart from irrigation, both blocks received identical treatment, and both were transplanted to 25 L pots at the same time.

The two treatments therefore differed in three respects at once: access to sensor data, the identity of the interpreter (a language model or a person) and the opportunity to irrigate (every 48 h in the LLM-guided block, at any time in the expert block). The trial can describe the combined outcome of these differences, but it cannot attribute that outcome to any one of them.



Figure 4. Layout of the trial: schematic of the expert and LLM-guided blocks on opposite halves of a single bench (pot positions were not randomised), and view of the pots in the greenhouse.

3.6. Query and Irrigation Protocol

Every 48 h, the operator inserted the probe into each pot of the LLM-guided block in turn. For each pot, a new conversation was opened for that plant, so that no conversation history was carried from one check to the next or from one plant to another; within a conversation, the earlier turns were passed back with each request as described in Section 3.4. The operator asked about the plant's water status in his own words. The questions were not scripted, and only the exchange reported in Section 4.4 was preserved. When the model recommended irrigation, the operator applied the midpoint of the volume range it suggested, measured with a graduated container. When the model did not recommend irrigation, no water was applied to that pot until the next check. No water was given to the LLM-guided block between checks.

3.7. Measurements and Derived Indicators

Water applied to each block was recorded as cumulative block totals, V, at two-week intervals up to 14 weeks after planting. Volumes for individual pots and the number and size of individual irrigation events were not retained. The relative difference in water applied at time t was calculated as

$$S(t) = 100 \times [V_{\mathrm{E}}(t) - V_{\mathrm{L}}(t)]/V_{\mathrm{E}}(t) \quad (2)$$

where V_{E} and V_{L} are the cumulative volumes applied to the expert and LLM-guided blocks. Because both quantities are single block totals, $S(t)$ is an arithmetic description of the difference between two series, not an estimate of a treatment effect. Plant height, from the soil surface to the apical meristem, was measured every 15 days up to day 105 and recorded as the mean of the six plants in each block; values for individual plants were not retained. At harvest, the tubers from each block were counted and their calibre (maximum diameter,

cm) was measured. Tuber fresh mass was not recorded and tubers were not graded for marketability, so tuber production is described only by number and maximum diameter, which are not measures of yield. Soil water content was measured only in the LLM-guided block, as part of the advisory protocol; it was not measured in the expert block, and soil aeration was not measured in either block. As a supporting descriptor, we also report the number of tubers harvested per litre of water applied,

$$WP_n = N/V \tag{3}$$

where N is the number of tubers harvested from a block and V is the cumulative water applied to it (tubers L^{-1}). WP_n is not water productivity in the conventional, fresh-mass sense and is not used to compare the treatments.

3.8. Data Analysis

Treatments were not replicated, and water, height and tuber data were recorded at block level. No inferential statistics were therefore possible, and none are presented. All results are block totals or block means for six pots and are reported descriptively, without measures of variability. Values quoted in the text for weeks and days were taken from the recorded series shown in Figures 5 and 6. In addition to the agronomic record, we examined the one advisory exchange that was preserved, recorded at crop emergence. Each element of the reply was compared with the context the model had received and with established agronomic reasoning for potato on Vertisols, so that we could see which parts of the advice followed from the data, which drew on general knowledge and which inherited errors from the context.

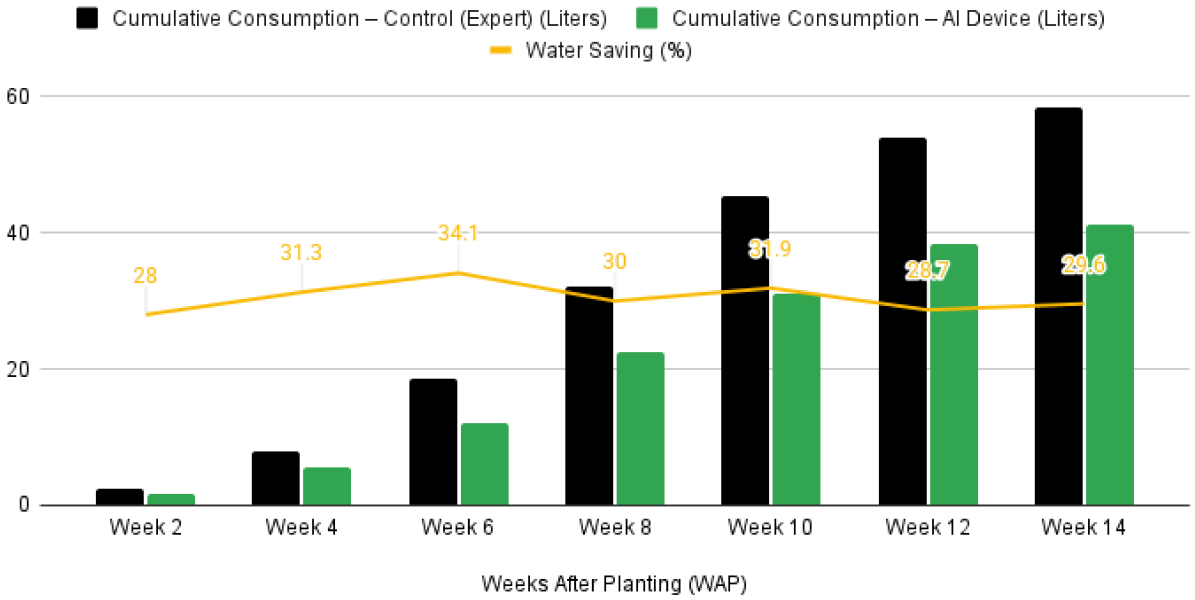


Figure 5. Cumulative water applied to the expert and LLM-guided blocks (bars) from week 2 to week 14 after planting, with the relative difference $S(t)$ (line, Equation (2)). Each bar is a block total for six pots; the treatments were not replicated. The line is an arithmetic description of the difference between the two series, not an inferential estimate. Legend labels in the figure: Control (Expert) = expert block; AI Device = LLM-guided block.

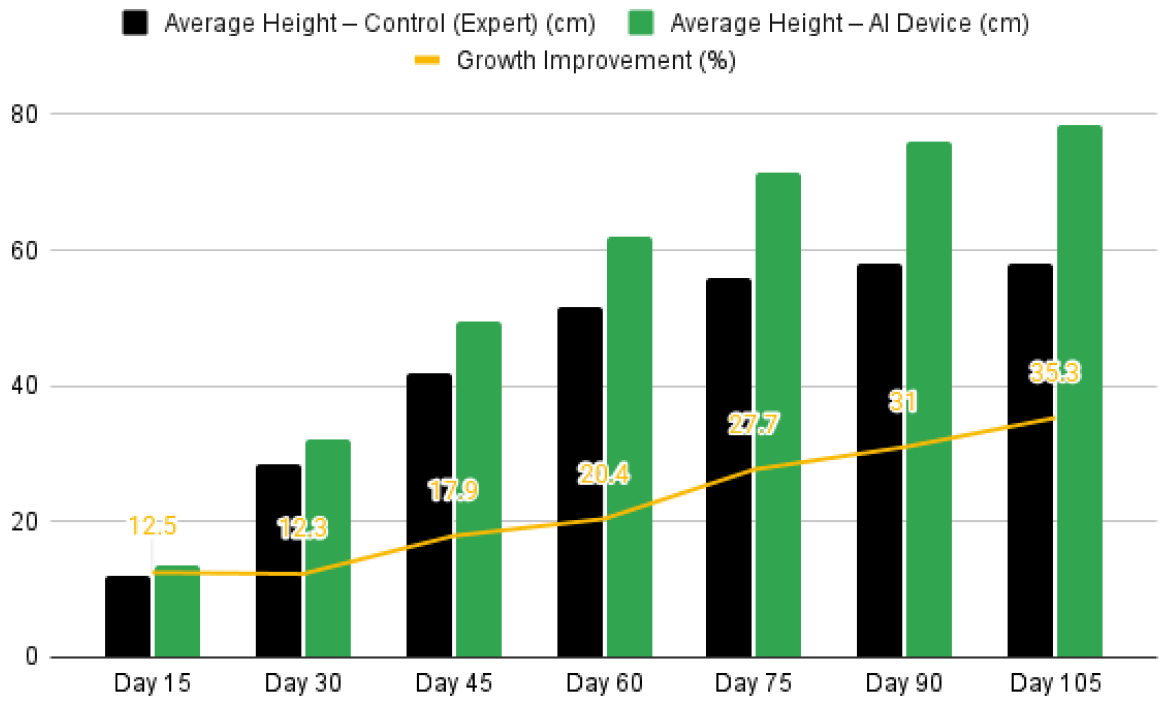


Figure 6. Mean plant height in the expert and LLM-guided blocks (bars) from day 15 to day 105, with the relative difference (line). Each bar is the mean of six plants in a single block; per-plant values were not retained, so no measure of variability is shown. The line is descriptive, not inferential. Legend labels in the figure: Control (Expert) = expert block; AI Device = LLM-guided block.

4. Results

4.1. Water Applied

The cumulative water applied to the two blocks diverged early in the season (Figure 5). At week 2, while the plants were still in 5 L pots, the LLM-guided block had received 28.0% less water than the expert block. The relative difference then stayed between 28% and 34% at every two-week record: 31.3% at week 4, 34.1% at week 6, 30.0% at week 8, 31.9% at week 10 and 28.7% at week 12. At week 14 the LLM-guided block had received 41.0 L and the expert block 58.2 L, a difference of 29.6%. In both blocks most of the water was applied between weeks 6 and 10, when the canopy expanded and tubers were initiated. Because individual irrigation events were not logged, these totals cannot show whether the difference arose from fewer events, from smaller volumes per event or from both. Because the expert block could be irrigated at any time and the LLM-guided block only at 48 h checks, part of the difference may reflect the schedule rather than the decisions made at each check.

4.2. Plant Growth

Mean plant height in the LLM-guided block was greater than in the expert block at every measurement (Figure 6). The difference was small in the first month, 12.5% at day 15 and 12.3% at day 30, and increased to 17.9% at day 45, 20.4% at day 60, 27.7% at day 75, 31.0% at day 90 and 35.3% at day 105. The shape of the two series also differed. Mean height in the expert block changed little after about day 75, levelling off at roughly 58 cm, whereas mean height in the LLM-guided block continued to increase through the bulking period and reached roughly 78 cm by day 105. These are means of six plants in a single block per treatment, and no measure of variability is available.

4.3. Tuber Number and Maximum Diameter

At harvest, the LLM-guided block produced 23 tubers and the expert block 14 (Table 2), or 3.8 and 2.3 tubers per plant. The tubers also differed in maximum diameter (Figure 7). Most tubers from the expert block had a maximum diameter between 1.0 and 4.5 cm, and the largest measured 6.5 cm. Most tubers from the LLM-guided block had a maximum diameter above 6.0 cm, and the largest measured 8.5 cm. Expressed per litre of water applied, the supporting descriptor WP_n was 0.24 tubers L^{-1} in the expert block and 0.56 tubers L^{-1} in the LLM-guided block. Tuber fresh mass and marketable grade were not recorded, so these figures describe tuber number and size only; they do not establish a difference in yield or in water productivity as conventionally defined.

Table 2. Block-level water applied and tuber record in the expert and LLM-guided blocks (six plants per block, one block per treatment).

Variable	Expert Block	LLM-Guided Block
Cumulative water at week 14 (L, block total)	58.2	41.0
Tubers harvested (n, block total)	14	23
Tubers per plant (block mean)	2.3	3.8
Maximum diameter of most tubers (cm)	1.0–4.5	>6.0
Largest tuber, maximum diameter (cm)	6.5	8.5
WP_n, supporting descriptor (tubers L^{-1})	0.24	0.56

Treatments were not replicated, so values are descriptive and no variability or test statistics can be given. Per-pot values, tuber fresh mass and marketable grade were not recorded. WP_n (Equation (3)) counts tubers per litre of water applied; it is not water productivity in the conventional fresh-mass sense.

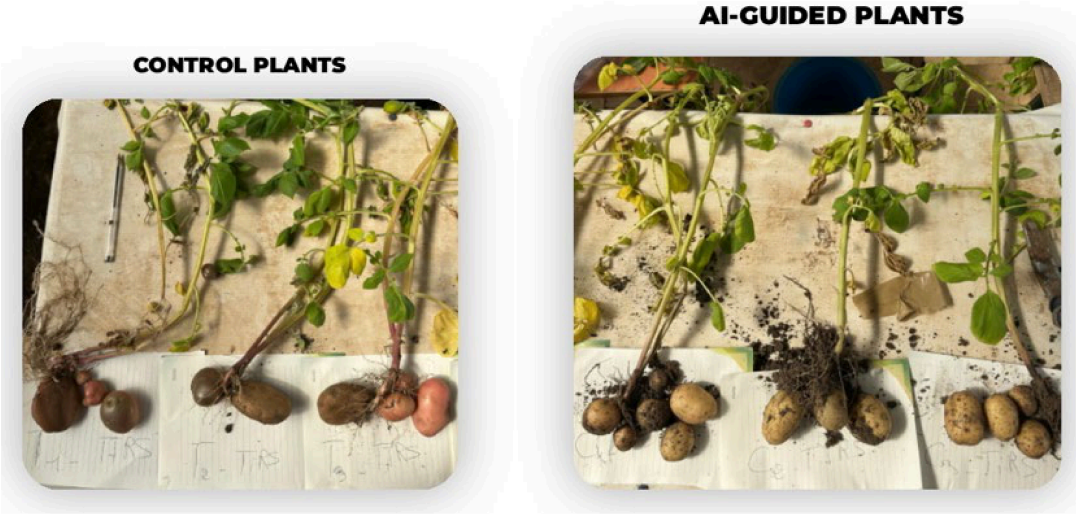


Figure 7. Tubers harvested from three plants of each block, photographed without selection: expert block (**left**) and LLM-guided block (**right**). The photograph has no scale bar; maximum diameters are reported in Table 2.

4.4. What the Dialogue Shows

The exchange below was recorded at the start of the trial, shortly after emergence, while the plants were in 5 L pots. The prompt contained the profile "Crop: *Solanum tuberosum*; Stage: germination; Soil: Clay/Tirs; Container: medium 5 L pot" and the reading "soil humidity 8% RH, temperature 29.4 °C". The operator asked, and the model replied (verbatim):

"How about water, how are we feeling?"

> "I am a Potato plant in the germination stage, and my soil temperature is currently 29.4 °C. However, my soil humidity is at 8% RH, which is critically low. Because I am growing in clay soil—which retains water well but requires careful saturation to avoid root asphyxia—I am feeling quite parched. To optimize my early root development, please provide approximately 250–500 mL of water to properly moisten the soil profile in this medium-sized pot. Thank you for your care!"

Table 3 breaks the reply into its components. Three features stand out. First, the reply drew on several parts of the context at once. It linked the low reading to the growth stage and framed its recommendation in terms of the soil type, warning against saturating a clay, which is the concern a soil scientist would raise for a Vertisol. This shows that the recommendation was generated in a context-dependent way. It does not show that the recommendation was agronomically correct: the suitability of the 250–500 mL volume was not validated independently, for example against a measured soil water deficit or plant water status, and the exchange provides no evidence on the accuracy of the advisor's irrigation recommendations. Second, every number in the reply can be traced to the input. The model did not invent readings, and it scaled its volume explicitly to the container described in the profile, which at this stage matched the 5 L pot the plant was growing in. The same dependence means that the volume would have been wrong had the container line not been updated when the plants were moved to 25 L pots; in this design, keeping it correct relied on the operator remembering to change it. Third, the model accepted the water reading under the label "% RH", a unit of air humidity, although the probe measures volumetric water content, and repeated the label in its reply without comment. Because the probe used the factory calibration, the 8% value itself is also uncertain in this soil (Section 3.3). The model treated it as exact. The exchange is a single example and cannot show how often such issues affected the advice. It does show which parts of the reply depend on the context, and therefore where errors in that context would propagate.

Table 3. Components of the advisory reply recorded at crop emergence, traced to the context the model received.

Element of the Reply	Source in the Context	Assessment
"germination stage"	Profile: growth stage	Correct; frames the advice around early root establishment
"soil temperature is currently 29.4 °C"	Sensor reading $D(t)$	Reported accurately
"8% ... critically low"	Sensor reading $D(t)$	Consistent with a dry clay at emergence, but the probe used its factory calibration, so the absolute value is uncertain in this soil
"clay soil ... careful saturation to avoid root asphyxia"	Profile: soil type, plus general knowledge	Draws on the stated soil type; the caution is consistent with Vertisol behaviour, but the advice was not independently validated
"250–500 mL ... in this medium-sized pot"	Profile: "medium 5 L pot"	Scaled to the stated container, which matched the 5 L pot in use at this stage; shows that recommended volumes depend directly on the profile; the suitability of the volume was not independently validated
"% RH"	Interface unit label	An air-humidity unit applied to soil water content; repeated rather than questioned

5. Discussion

5.1. What the Case Record Shows, and What It Does Not

Over a full potato cycle, the block managed through the advisor received less water than the block managed by the technician, and its plants grew taller and set more tubers of larger maximum diameter. This is the record of one season, with one block per treatment on one bench, and it supports three limited statements. The advisor can be operated over a whole crop cycle with inexpensive hardware and a general-purpose model. Following its advice did not lead to visible failure of the crop. And its recommendations were, on the whole, more sparing with water than the technician's practice. The record cannot show that the advisor caused the differences in growth or tuber set. Bench position, access to sensor data, the 48 h irrigation schedule and the identity of the decision-maker all changed together between the blocks (Section 3.5), and without replication there is no way to estimate how much of the observed difference any of them would account for.

One explanation that the record does not exclude, and that deserves a direct test, is that the expert block was over-irrigated. A Vertisol surface can crack and look dry while the clay below remains moist [38], and excess water cannot drain from a pot, so judging by eye could lead to watering a root zone that was already wet, with a possible loss of aeration [40,41]. This mechanism was not observed. Soil water content was not measured in the expert block and aeration was not measured in either block, so it is offered as a hypothesis for a future trial, not as an explanation of the results. It could be tested by inserting the same probe in the expert pots with the readings withheld from the technician, by weighing pots and collecting drainage, or by measuring stomatal conductance or canopy temperature.

Tuber number and maximum diameter are not substitutes for yield. A greater number of tubers is not necessarily agronomically better, and water productivity and marketability in potato are conventionally expressed on a fresh-mass basis with tubers graded by size [44]. Since neither mass nor grade was recorded, we draw no conclusion about yield, and WP_n is reported only as a supporting descriptor.

5.2. What the Language Model May Contribute

The design does not separate the contribution of the language model from that of the sensor. The same reading could have driven a fixed threshold rule of the kind used by commercial soil moisture controllers [12,13], or could have been handed to the technician. Either control might have produced similar water use. What the single preserved exchange does show is the kind of contribution a language model can make in principle. It interpreted the reading in terms of the crop's stage and the soil's properties, without thresholds in the prompt and without fine-tuning, and explained its advice in plain language. This is evidence of context-dependent recommendation generation, not of agronomic accuracy, since the volumes recommended were not validated independently. Whether that interpretation improves decisions over a well-chosen threshold is an empirical question that this study cannot answer. The water difference we recorded is of the same order as the 37.5% reported by Saranya et al. [21] for a rule-constrained LLM advisor over one week, but neither study isolates the language model's contribution.

The operator remained part of the loop throughout, and we regard this as a strength of the design rather than a stage to be automated away. Tzachor et al. [18] and Mourtzinis et al. [19] both concluded that LLM advice becomes trustworthy when a person is in a position to check it. A conversational interface makes that checking possible, because the operator can ask why, ask again, or ask about something the designer did not anticipate. The present study did not assess how operators used that possibility.

5.3. Letting the Crop Speak: Persona and Trust

We gave the plant a first-person voice in the expectation that it would make the advice easier for non-specialists to engage with. Studies outside agriculture show that people engage more readily with, and ascribe more competence to, systems that present in human-like ways [36,37]. We did not measure engagement, comprehension or trust in this study, so these remain potential benefits supported by that literature rather than findings. The same literature points to a risk. A persuasive voice invites the kind of reliance in which people stop checking automated advice [35], and the dialogue in Section 4.4 shows that the voice was equally confident when part of its input was mislabelled. We therefore suggest that a speaking crop should also show its evidence. Each reply should display the reading and profile values it used, together with their units, so that an operator can see at a glance that a soil water reading has been reported as air humidity, or that the container described is not the one the plant is in.

5.4. Context Fidelity and Other Failure Modes

The preserved exchange suggests that, once an LLM is grounded in live data, the main risk is not that it invents information. It is that it reasons faithfully from a context that is wrong, incomplete or out of date. Two examples arose in this study. The moisture reading was passed to the model under an air-humidity unit, and the model accepted and repeated it. The container line, correct at emergence, had to be changed by hand when the plants were transplanted; had it not been, every volume would have been scaled to a container one-fifth the actual size. Table 4 lists these and other failure modes that are documented for comparable systems, with a safeguard for each. Several of the safeguards are simple. A numeric, validated container field shown with each reply would make an out-of-date profile visible. A fixed unit schema passed with every reading would have prevented the unit error. Writing the allowable depletion and target water range for the crop [23] into the prompt, as Chen et al. [22] did with hydraulic limits, would give the model limits against which to check its own volumes. Retrieval from local extension guidelines [32] would anchor its general knowledge in regional practice. Pinning a model snapshot, fixing the sampling settings and logging every request and response would make each recommendation reproducible and auditable [47]. None of these safeguards requires a better model. All of them require more careful engineering of the context the model receives and of the record that it leaves.

Table 4. Failure modes of a sensor-grounded conversational irrigation advisor and proposed safeguards.

Failure Mode	Evidence	Consequence	Safeguard
Unit or label error	This study: soil water content labelled “% RH”	Misread thresholds; loss of operator trust	Fixed unit schema sent with every reading
Out-of-date profile	This study: container line had to be changed by hand when plants moved from 5 L to 25 L pots	Volumes scaled to the wrong container if not updated	Numeric, validated profile fields; profile shown with each reply
Sensor calibration error	This study: factory calibration in a clay-rich soil [11,45]	Systematic bias in all advice	Soil-specific calibration; plausibility checks on readings

Table 4. *Cont.*

Failure Mode	Evidence	Consequence	Safeguard
Unsupported or numerically insensitive output	Ji et al. [34]; Chen et al. [22]	Volumes not matched to the actual deficit	Crop and soil water limits in the prompt; range check on every recommended volume
Non-reproducible output	This study: no pinned snapshot, default sampling, no logs [47]	Recommendations cannot be audited or repeated	Pinned snapshot, fixed sampling settings, full request and response log
Stale growth stage	Operator-updated profile (this design)	Advice inappropriate to the current stage	Stage proposed automatically from days after planting or thermal time
Over-trust in the persona	Parasuraman and Manzey [35]	Errors go unchallenged	Display evidence and uncertainty; keep the final decision with the operator

5.5. Limitations

The limitations of this study are substantial, and they define how its results should be read. Each treatment consisted of six pots grouped on one half of a single bench, without randomisation or rotation, so bench position is confounded with treatment. Water was logged as block totals, and height as block means, so the twelve pots yield two values per variable, and no measure of variability or inferential test is possible. The treatments also differed in access to sensor data, in the identity of the decision-maker and in the opportunity to irrigate, so the trial measures the net outcome of the sensor-plus-advisor package and cannot separate their contributions. Soil water content was not measured in the expert block, and aeration was not measured at all. Tuber fresh mass and marketable grade were not recorded. The probe was used with its factory calibration in a clay-rich soil, so absolute water content values are uncertain [45]. The model snapshot and sampling settings were not fixed and responses were not logged, so recommendations cannot be reproduced or independently verified, their consistency across the season cannot be assessed, and only one exchange was preserved for analysis. The plants were grown in 5 L pots for the first three to four weeks and in 25 L pots thereafter, so the early water record reflects the smaller containers. The comparison was with one technician, the operator who applied the advice knew which block was which, and the trial ran for a single season under greenhouse conditions with no laboratory characterisation of the substrate. Finally, GPT-5 is a proprietary model whose behaviour can change between versions [46]. For these reasons we present the study as a proof of concept and a case record. Its results indicate that the approach merits a properly designed trial, not that it outperforms expert management.

5.6. Towards a Replicated Trial and Field Deployment

The limitations above translate directly into the design of the next trial. Pots or plots should be the experimental units, arranged in randomised complete blocks along the bench and rotated at least weekly. To separate the contribution of the sensor from that of the interpreter, the trial should include at least four arms: a technician without sensor data, the same technician given the probe readings, a fixed volumetric water content threshold applied to the same probe, and the language-model advisor. The same probe should log soil water in every pot, with readings withheld from the arms that are not meant to see them, and pot weight and drainage should be recorded so that the over-irrigation hypothesis can be tested. Irrigation should be logged per pot and per event, and harvest should record tuber fresh mass and marketable grade per pot. The probe should be calibrated against

gravimetric samples of the trial soil, with the calibration equation, RMSE and R^2 reported and the working range stated. The model snapshot and sampling settings should be pinned and every request and response logged, and the questions put to the advisor should follow a fixed script. The trial should run over more than one season.

Beyond the greenhouse, probes at several depths, combined with reference evapotranspiration from a nearby weather station, would give the model a view of the whole root zone and of atmospheric demand [23]. Open-weight models running on a local gateway would remove the dependence on connectivity and on a single commercial provider. For Moroccan smallholders, a voice interface in Darija, Tamazight or French may matter more than any gain in model accuracy, because the value of the system lies in reaching growers whom dashboards do not reach [10]. Connecting the advisor to valves is technically straightforward, but we would take that step only behind hard limits and with the operator retaining the final say.

6. Conclusions

The principal contribution of this proof-of-concept study is a demonstration of technical feasibility. Real-time readings from a low-cost soil probe and a short description of the crop can be integrated with a general-purpose language model in a conversational advisory interface, and the system can be operated throughout a complete potato cycle on a difficult clay soil, using inexpensive hardware and no fine-tuning. The study does not establish the agronomic effectiveness of the resulting irrigation recommendations. In a single unreplicated greenhouse trial, the block irrigated on the advisor's recommendation received less water than a block managed by an experienced technician without sensor data, and its plants grew taller and set more tubers of larger maximum diameter. These observations are descriptive. The design does not allow them to be attributed to the language model rather than to the sensor, bench position or irrigation schedule, and yield was not measured. The recorded dialogue showed that the advisor's main vulnerability is its fidelity to its inputs: it reasons faithfully from whatever context it is given, including mislabelled units and profile details that must be kept up to date by hand. Letting the crop speak therefore requires the context to be engineered as carefully as the sensor itself. A randomised, replicated trial that includes threshold-based and sensor-informed human controls, per-pot measurements of water, soil moisture and tuber mass, and a pinned, fully logged model is the necessary next step.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriengineering8110468/s1, File S1: Full system message sent to GPT-5 with each advisory request.

Author Contributions: Conceptualization, I.E. and A.T.; methodology, I.E., A.M., D.B. and Y.A.; software, I.E.; validation, I.E., A.M. and Y.A.; formal analysis, I.E.; investigation, A.M. and G.G.; resources, I.E., A.M., D.B., Y.A. and G.G.; data curation, Y.A.; writing—original draft preparation, I.E.; writing—review and editing, I.E. and O.H.; visualization, I.E. and A.T.; supervision, O.H.; project administration, O.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Center for Scientific and Technical Research (CNRST) under the "PhD-Associate Scholarship–PASS" program, within the framework of the National Strategy "Morocco Digital 2030" (Horizon 2027).

Data Availability Statement: All data recorded in this study are reported in the article: block-level cumulative water applied, mean plant height, and tuber number and maximum diameter. Values for individual pots, individual irrigation events and model responses other than the exchange reported in Section 4.4 were not retained. The full system message is provided as Supplementary File S1.

Acknowledgments: The authors thank the technical staff of the Ibn Tofail University greenhouse for their help throughout the experiment. GPT-5 (OpenAI, San Francisco, CA, USA) is a component of the system studied here, and its use is described in Section 3.4. During the preparation of this manuscript, the authors also used AI-assisted tools for the purposes of improving the language and structure of the text. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application programming interface
CNRST	National Center for Scientific and Technical Research
FAO	Food and Agriculture Organization of the United Nations
IoT	Internet of Things
LLM	Large language model
ML	Machine learning
RH	Relative humidity
VWC	Volumetric water content
WP_n	Count-based water productivity (tubers L^{-1})

References

1. FAO. *The State of the World's Land and Water Resources for Food and Agriculture 2021—Systems at Breaking Point*; FAO: Rome, Italy, 2022. [CrossRef]
2. Gumus, V.; El Moçayd, N.; Seker, M.; Seaid, M. Future projection of droughts in Morocco and potential impact on agriculture. *J. Environ. Manag.* **2024**, *367*, 122019. [CrossRef] [PubMed]
3. Agripedia. Pomme de Terre. Available online: https://agripedia.ma/fr/maraichage/pomme-de-terre (accessed on 30 September 2026). (In French)
4. Obidiegwu, J.E.; Bryan, G.J.; Jones, H.G.; Prashar, A. Coping with drought: Stress and adaptive responses in potato and perspectives for improvement. *Front. Plant Sci.* **2015**, *6*, 542. [CrossRef] [PubMed]
5. Shock, C.C.; Wang, F.-X. Soil water tension, a powerful measurement for productivity and stewardship. *HortScience* **2011**, *46*, 178–185. [CrossRef]
6. Tzounis, A.; Katsoulas, N.; Bartzanas, T.; Kittas, C. Internet of Things in agriculture, recent advances and future challenges. *Biosyst. Eng.* **2017**, *164*, 31–48. [CrossRef]
7. Ray, P.P. Internet of things for smart agriculture: Technologies, practices and future direction. *J. Ambient Intell. Smart Environ.* **2017**, *9*, 395–420. [CrossRef]
8. Elijah, O.; Rahman, T.A.; Orikumhi, I.; Leow, C.Y.; Hindia, M.N. An overview of Internet of Things (IoT) and data analytics in agriculture: Benefits and challenges. *IEEE Internet Things J.* **2018**, *5*, 3758–3773. [CrossRef]
9. Levidow, L.; Zaccaria, D.; Maia, R.; Vivas, E.; Todorovic, M.; Scardigno, A. Improving water-efficient irrigation: Prospects and difficulties of innovative practices. *Agric. Water Manag.* **2014**, *146*, 84–94. [CrossRef]
10. Shang, L.; Heckelei, T.; Gerullis, M.K.; Börner, J.; Rasch, S. Adoption and diffusion of digital farming technologies—Integrating farm-level evidence and system interaction. *Agric. Syst.* **2021**, *190*, 103074. [CrossRef]
11. Abdelmoneim, A.A.; Al Kalaany, C.M.; Khadra, R.; Derardja, B.; Dragonetti, G. Calibration of low-cost capacitive soil moisture sensors for irrigation management applications. *Sensors* **2025**, *25*, 343. [CrossRef] [PubMed]
12. Cardenas-Lailhacar, B.; Dukes, M.D. Precision of soil moisture sensor irrigation controllers under field conditions. *Agric. Water Manag.* **2010**, *97*, 666–672. [CrossRef]
13. Romero, R.; Muriel, J.L.; García, I.; Muñoz de la Peña, D. Research on automatic irrigation control: State of the art and recent results. *Agric. Water Manag.* **2012**, *114*, 59–66. [CrossRef]
14. Goap, A.; Sharma, D.; Shukla, A.K.; Rama Krishna, C. An IoT based smart irrigation management system using machine learning and open source technologies. *Comput. Electron. Agric.* **2018**, *155*, 41–49. [CrossRef]
15. Vij, A.; Vijendra, S.; Jain, A.; Bajaj, S.; Bassi, A.; Sharma, A. IoT and machine learning approaches for automation of farm irrigation system. *Procedia Comput. Sci.* **2020**, *167*, 1250–1257. [CrossRef]

16. Benos, L.; Tagarakis, A.C.; Dolias, G.; Berruto, R.; Kateris, D.; Bochtis, D. Machine learning in agriculture: A comprehensive updated review. *Sensors* **2021**, *21*, 3758. [CrossRef] [PubMed]
17. Silva, B.; Nunes, L.; Estevão, R.; Aski, V.; Chandra, R. GPT-4 as an agronomist assistant? Answering agriculture exams using large language models. *arXiv* **2023**, arXiv:2310.06225. [CrossRef]
18. Tzachor, A.; Devare, M.; Richards, C.; Pypers, P.; Ghosh, A.; Koo, J.; Johal, S.; King, B. Large language models and agricultural extension services. *Nat. Food* **2023**, *4*, 941–948. [CrossRef] [PubMed]
19. Mourtzinis, S.; Silva, T.S.; Lo, J.C.M.; Smith, D.L.; Conley, S.P. A human in the loop approach to applying large language models for farm management insight. *Sci. Rep.* **2025**, *16*, 1273. [CrossRef] [PubMed]
20. An, T.; Zhou, Y.; Zou, H.; Yang, J. IoT-LLM: A framework for enhancing large language model reasoning from real-world sensor data. *Patterns* **2026**, *7*, 101429. [CrossRef] [PubMed]
21. Saranya, S.N.; Madhumala, R.B.; Samuel, S.D.; Yaplar, S.G.; Shashwath, V.; Paul, V.S. AgriBot: An LLM-guided IoT system for real-time soil monitoring and irrigation advice. In *Proceedings of the 2026 Sixth International Conference on Advances in Electrical, Computing, Communications and Sustainable Technologies (ICAECT)*; IEEE: Piscataway, NJ, USA, 2026; pp. 1–6. [CrossRef]
22. Chen, J.; Wang, H.; Xu, J.; Zhang, B.; Zhang, H.; Ye, Z.; Zhang, Z.; Liu, X.; Chen, R. Closed-loop autonomous scheduling of multi-level irrigation canal-gate systems using an LLM-agent framework: Prompting strategies and model heterogeneity. *Comput. Electron. Agric.* **2027**, *256*, 112444. [CrossRef]
23. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*; FAO Irrigation and Drainage Paper 56; FAO: Rome, Italy, 1998.
24. Thompson, R.B.; Gallardo, M.; Valdez, L.C.; Fernández, M.D. Using plant water status to define threshold values for irrigation management of vegetable crops using soil moisture sensors. *Agric. Water Manag.* **2007**, *88*, 147–158. [CrossRef]
25. Muñoz-Carpena, R.; Li, Y.C.; Klassen, W.; Dukes, M.D. Field comparison of tensiometer and granular matrix sensor automatic drip irrigation on tomato. *HortTechnology* **2005**, *15*, 584–590. [CrossRef]
26. Kamienski, C.; Soininen, J.-P.; Taumberger, M.; Dantas, R.; Toscano, A.; Salmon Cinotti, T.; Filev Maia, R.; Torre Neto, A. Smart water management platform: IoT-based precision irrigation for agriculture. *Sensors* **2019**, *19*, 276. [CrossRef] [PubMed]
27. Pylianidis, C.; Osinga, S.; Athanasiadis, I.N. Introducing digital twins to agriculture. *Comput. Electron. Agric.* **2021**, *184*, 105942. [CrossRef]
28. Verdouw, C.; Tekinerdogan, B.; Beulens, A.; Wolfert, S. Digital twins in smart farming. *Agric. Syst.* **2021**, *189*, 103046. [CrossRef]
29. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022. [CrossRef]
30. Zhao, B.; Jin, W.; Del Ser, J.; Yang, G. ChatAgri: Exploring potentials of ChatGPT on cross-linguistic agricultural text classification. *Neurocomputing* **2023**, *557*, 126708. [CrossRef]
31. Zhu, H.; Qin, S.; Su, M.; Lin, C.; Li, A.; Gao, J. Harnessing large vision and language models in agriculture: A review. *Front. Plant Sci.* **2025**, *16*, 1579355. [CrossRef] [PubMed]
32. Lewis, P.; Perez, E.; Piktus, A.; Petroni, F.; Karpukhin, V.; Goyal, N.; Küttler, H.; Lewis, M.; Yih, W.; Rocktäschel, T.; et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 9459–9474. [CrossRef]
33. Yang, B.; Zhang, Y.; Chen, Y.; Feng, L.; Xu, X.; Aierken, N.; Li, S. AgriAgent: Contract-driven planning and capability-aware tool orchestration in real-world agriculture. *arXiv* **2026**, arXiv:2601.08308. [CrossRef]
34. Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y.J.; Madotto, A.; Fung, P. Survey of hallucination in natural language generation. *ACM Comput. Surv.* **2023**, *55*, 248. [CrossRef]
35. Parasuraman, R.; Manzey, D.H. Complacency and bias in human use of automation: An attentional integration. *Hum. Factors* **2010**, *52*, 381–410. [CrossRef] [PubMed]
36. Epley, N.; Waytz, A.; Cacioppo, J.T. On seeing human: A three-factor theory of anthropomorphism. *Psychol. Rev.* **2007**, *114*, 864–886. [CrossRef] [PubMed]
37. Blut, M.; Wang, C.; Wünderlich, N.V.; Brock, C. Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *J. Acad. Mark. Sci.* **2021**, *49*, 632–658. [CrossRef]
38. Stewart, R.D.; Abou Najm, M.R.; Rupp, D.E.; Selker, J.S. Modeling multidomain hydraulic properties of shrink-swell soils. *Water Resour. Res.* **2016**, *52*, 7911–7930. [CrossRef]
39. Moussadek, R.; Mrabet, R.; Zante, P.; Lamachère, J.M.; Pépin, Y.; Le Bissonnais, Y.; Ye, L.; Verdoodt, A.; Van Ranst, E. Impact of tillage and residue management on the soil properties and water erosion of a Mediterranean Vertisol. *Can. J. Soil Sci.* **2011**, *91*, 627–635. [CrossRef]
40. da Silva, A.P.; Kay, B.D.; Perfect, E. Characterization of the least limiting water range of soils. *Soil Sci. Soc. Am. J.* **1994**, *58*, 1775–1781. [CrossRef]
41. Jackson, M.B.; Colmer, T.D. Response and adaptation by plants to flooding stress. *Ann. Bot.* **2005**, *96*, 501–505. [CrossRef] [PubMed]

42. Shock, C.C.; Feibert, E.B.G.; Saunders, L.D. Potato yield and quality response to deficit irrigation. *HortScience* **1998**, *33*, 655–659. [CrossRef]
43. Mora-Sanhueza, R.; Tighe-Neira, R.; López-Olivari, R.; Inostroza-Blancheteau, C. Assessment of different irrigation thresholds to optimize the water use efficiency and yield of potato (*Solanum tuberosum* L.) under field conditions. *Plants* **2025**, *14*, 1734. [CrossRef] [PubMed]
44. Niu, Y.; Wang, L.; Luo, Z.; Fudjoe, S.K.; Palta, J.A.; Li, L.; Li, S. Effects of irrigation practices on potato yield and water productivity: A global meta-analysis. *Agronomy* **2025**, *15*, 1942. [CrossRef]
45. Kizito, F.; Campbell, C.S.; Campbell, G.S.; Cobos, D.R.; Teare, B.L.; Carter, B.; Hopmans, J.W. Frequency, electrical conductivity and temperature analysis of a low-cost capacitance soil moisture sensor. *J. Hydrol.* **2008**, *352*, 367–378. [CrossRef]
46. Singh, A.; Fry, A.; Perelman, A.; Tart, A.; Ganesh, A.; El-Kishky, A.; McLaughlin, A.; Low, A.; Ostrow, A.J.; Ananthram, A.; et al. OpenAI GPT-5 system card. *arXiv* **2026**, arXiv:2601.03267. [CrossRef]
47. Ouyang, S.; Zhang, J.M.; Harman, M.; Wang, M. An empirical study of the non-determinism of ChatGPT in code generation. *ACM Trans. Softw. Eng. Methodol.* **2025**, *34*, 42. [CrossRef]
48. Brien, C.J.; Berger, B.; Rabie, H.; Tester, M. Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems. *Plant Methods* **2013**, *9*, 5. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.