Institute for Physical AI @ Bailey Military Institute · The Glass Lab & The Charlot Lab

SENSING · AGRICULTURE · MEASUREMENT

Technical Report TR–2026–44
Research / Position · Preprint v1.0
26 August 2026

SENSING · AGRICULTURE · MEASUREMENT

The Unknown Fraction

Agricultural sensing reports what it measured. It does not report what it could not see. Three instruments over three unrelated physics find the same division, and four industries already pay for the quantity none of them computes.

Ronnie Gomez, Industrial Research Fellow · The Charlot Lab
David Jean Charlot, PhD, Dean of Physical AI · The Charlot Lab
Charles R. Glass, EdD, Executive Director, Bailey Military Institute · The Glass Lab

Correspondence: contact@physicalai-bmi.org · physicalai-bmi.org

Instruments: the grain bin · the soil core · the canopy

3 instruments, each runnable in a browser · 3 unrelated physics, one convergent division · 6 quantities recorded as not located · 1 of our own claims withdrawn

Abstract. Agriculture is instrumented at two extremes. Satellites resolve the canopy surface at ten metres; a point probe resolves a few cubic centimetres. Every agronomic decision is made in the volume between them, and this review did not locate an agricultural sensing system that reports what fraction of that volume it cannot see. We define the unknown fraction as a property of sensing geometry rather than of a model, compute it in three settings whose physics have nothing in common, and find the same division each time: the field average is well resolved and the localized answer is not. We show that four industries already carry this quantity as a cost under four different names, that a fifth name for it has been published in every variogram fitted for decades, and that the reach of a modality, not the count of its sensors, is the variable that moves it. One claim we published during

this work was wrong; it is withdrawn in §10 and the corrected version is stronger.

1. What is in scope, and what does this continue?

This report is about what an agricultural sensing system can and cannot resolve, treated as a geometric property of where its sensors are and how far each one reaches. It is not about which sensor to buy, and it does not propose hardware.

It continues three lines of Institute work. *OmniSense* (TR-2026-08) treats a world model not as internal state read from onboard sensors but as a field defined over a volume by a mesh of nodes, some carried by the system, some on peers, some anchored in the space, resolving that volume into occupied, confirmed empty, and unknown.[1] This report imports that framing and asks what it means for a farm. *Physical AI and Logistics Opportunities in Agriculture* (TR-2026-32) sizes the agricultural estate and its labour market.[2] That is a different question from this one: it asks what can be done economically, and this asks what can be known. *The Human Layer* (TR-2026-18) argues that in autonomous systems the binding constraint is the human and the pipeline that produces the competence.[3] §11 extends that argument with one competency agriculture forces and the airspace work did not.

Three instruments accompany this report. Each runs in a browser, each states its model in its own text, and each is linked from the header above. Every figure they compute is reproduced here.

2. What is the unknown fraction?

Take a volume in which decisions are made: a grain mass, a root zone, a tree canopy, a pen of fish. Place sensors in or around it. Each sensor reaches some distance, set by the physics of whatever carries its signal. The union of those reaches is the resolved volume. Everything else is the unknown fraction.

The definition is deliberately geometric. It does not ask whether a model can predict the unresolved part; it asks whether any sensor touched it. That distinction is the whole content of this report, because the two answers differ by orders of magnitude and only one of them is reported.

The quantity has a second reading that makes it operational. If an event begins at a point drawn uniformly from the volume, the probability that no sensor reaches it is exactly the unknown fraction. The fraction of the volume you cannot see and the probability you miss what happens in it are the same number.

3. Why is agriculture the hard case?

In a warehouse the barrier that blocks sensing is a wall: static, dry, and not the subject of the measurement. In a field it is biomass and soil, which are dielectric, water bearing, seasonally changing, and themselves the quantity of agronomic interest.

In agriculture the barrier and the signal are the same substance. Soil moisture attenuates the radio that would measure soil moisture, and the attenuation is worst in the moist and clay rich soils most worth reading.[4] Canopy density occludes the fruit whose count is set by canopy density. Grain mass insulates the hot spot that grain mass creates. The occlusion is not noise on the measurement. It is the measurement, inverted.

That is why agriculture is the hard case for embodied perception rather than an easy application of it, and it is why the unknown fraction is worth computing here first. Four barriers appear across the sectors we surveyed, and each fails differently, so each needs a different engineering change.

Barrier	What fails	What moves it
Soil	Attenuation rises with the moisture being measured; ground penetrating radar reaches under 10 m in most ground, under 1 m in conductive clays, and centimetres in moist high conductivity soil[5]	Magnetic induction or acoustic carriers; buried node to surface relay geometry

Canopy	Occlusion sets a ceiling no resolution fixes	Multi view geometry; semantic scene completion over unobserved space
Vegetation over soil	Radar backscatter over annual crops is dominated by canopy structure, so retrieval needs correction[6]	In situ nodes as the boundary condition that constrains the inversion
Power	Sampling density is capped by joules; naive schedules drain a node in days[7]	Soil microbial fuel cells reported at 68 times the power a buried node needs[8]

4. What does a grain bin actually see?

We begin where the physics is cleanest. A grain bin is rigid, bounded, and already monitored, and the standard instrument is a temperature cable carrying probes at intervals.

Heat moves through grain by diffusion, so a probe senses a sphere whose radius is the diffusion length $L = \sqrt{(\alpha t)}$. At a grain thermal diffusivity near 1.1×10^{-7} m²/s, a probe reaches **0.26 m in a week**. Thirty probes on five cables in a 9.1 by 12.2 m bin therefore reach 2.2 of its 793 cubic metres.

Quantity	Value
Volume monitored	0.27%
Unknown fraction	99.73%
Probability a hot spot at a random point goes unseen	99.73%, the same number
Probes needed to reach every cubic metre	about 11,000

The last row prices the obvious response out of reach, and points at the correct one, which is not more probes. Sensor count to fill a volume scales as $1/r^3$, so reach is worth far more than count.

Carbon dioxide transport through a grain bulk is also diffusion, which makes the comparison exact rather than rhetorical. Measured effective CO_2 diffusivity through bulk corn is 3.10 to 3.93×10^{-6} m²/s across 10 to 30 °C and 14.0 to 18.8% moisture.[9] Both reaches are √(kt) and differ only by transport coefficient, so **CO_2 reaches 5.6 times further at every time horizon and needs 179 times fewer nodes**. The ratio is fixed by √(D/α) and does not move when the operator changes how long they are willing to wait. For the same bin at seven days: 11,039 temperature probes, or 62 CO_2 intakes.

This corrects an attribution. Trade literature describes CO_2 as a more sensitive spoilage indicator than temperature.[10] Sensitivity is not the mechanism. The gas simply gets further, so each sensor stands for far more grain. An industry practice justified by chemistry turns out to be justified by geometry, and the geometry says how much better it is.

5. What does a soil core know?

The second setting is the one carrying the largest budget. Soil sampling consumes 25 to 40% of a soil carbon project, field collection runs $15,000 to $30,000 for a medium farm, grid sampling needs about one technician per 100 hectares, and laboratory turnaround is four to eight weeks.[11] The sampling design is the single largest cost decision in the work.

A 100 hectare project taking 20 cores per hectare to 30 cm physically samples **0.63 parts per million** of the soil it will make a claim about, for roughly $24,000 in cores. On its own that figure is arresting and unfair, and the instrument says so. Soil carbon is spatially autocorrelated, so a core speaks for far more ground than it occupies.

The correction turns out to invert the picture. Published autocorrelation ranges for soil organic carbon span roughly 234 m to 2,102 m across sites.[12] At a 234 m range a single core speaks for 17 hectares, so six cores would cover a 100 hectare field areally, and the project took two thousand. Every published range leaves these designs **over sampled areally by hundreds of times**. For the field mean, they are not short of cores.

What they are short of is anything below the sampling spacing, and that quantity already has a name, a number, and a publication history. The **nugget** is the share of variance at separations shorter than any pair of samples. Reported nugget to sill ratios of 0.20 and 0.31 are both described as strong spatial dependence.[12] The nugget does not fall when cores are added at the same spacing; only sampling at shorter separations can estimate it at all.

So a fifth to nearly a third of the variance is unresolved by construction and is conventionally read as noise. It is not noise. It is unresolved signal, and it is where the contested questions live: additionality and loss events ask whether something happened *somewhere*, not what the field averages. **Geostatistics has been publishing the unknown fraction for decades under a name that made it sound like an error term.**

One further exposure follows from the range. Across the published span the areal requirement moves by roughly eighty times, and the sampling design does not change, because the range is assumed rather than measured for the field in hand. This review did not locate a protocol requirement that a project measure the autocorrelation range of the field it is sampling. The confidence rests on a parameter the project never bought.

6. What does a camera miss in a canopy?

The third setting is geometric rather than diffusive or statistical. Machine vision cannot count fruit the canopy hides, so yield estimation counts what it sees and multiplies by a correction factor calibrated on hand counted sample trees.

That factor has been measured, and the measurement is the finding. Single side visibility across real apple orchards runs **40.85% to 79.83%**, varying with trimming and canopy density, so the correction factor those orchards imply runs 1.25 to 2.45.[13] The published objection to hand calibration is not that the arithmetic fails but that selecting a representative tree is problematic given the variation in canopy density between trees in any orchard.[14]

Foliage extinguishes a sight path the way it extinguishes light, so the visible fraction is the Beer–Lambert transmittance in the leaf area between camera and fruit, integrated over fruit depth. With extinction coefficients in their published range of about 0.3 for erect foliage to 1.0

for horizontally held leaves,[15] the model spans the measured 41 to 80% band at k·LAI between about 0.5 and 2.2. At LAI 4.0 and k 0.55 it returns 40% visible and a 2.47 correction, against a measured dense end of 40.85% and 2.45. That agreement from independent parameters is the only reason the model is worth using.

Orchards are discontinuous canopies, where interception reaches an asymptote well short of what a continuous Beer–Lambert canopy would give,[16] so we treat these numbers as the shape of the problem rather than a calibrated orchard model.

7. Why do three unrelated physics converge?

The three settings share no mechanism. One is thermal diffusion, one is spatial statistics, one is geometric occlusion. They were built in that order and the third was not expected to agree with the first two. Each separates the same pair of questions.

Setting	Barrier	The mean question	The local question
Grain bin	heat diffusion	±0.8% on the field average	99.7% chance of missing the hot spot
Soil core	spatial statistics	over covered by 333 times	20 to 31% nugget, unreachable
Orchard	canopy occlusion	±0.8% on the block total	±22% on any individual tree

The reason for the convergence is that interpolation covers the gaps between sensors when the quantity varies smoothly, and does not when it does not. A field average is an integral over a smooth field, and sparse sampling estimates integrals well. A spoilage pocket, a disease focus, an insect colony, an unapplied practice and a leak are localized events, and there is nothing to interpolate them from.

This also answers the strongest objection to the whole line of work. Placement studies find five well sited soil probes reach about 2% estimation error and ten reach about 1%, with little gained afterwards.[17] That result is correct and it does not conflict with anything here. It is an answer to the mean question. **A sensing system**

that reports only the mean answer is not wrong. It is answering the easier question and not saying which one it answered.

Every decision that acts on a place rather than on an average sits in the second column: thinning, irrigation scheduling, harvest timing, targeted treatment, additionality, and the decision to unload a bin.

8. Who already pays for this?

The unknown fraction is not a new metric in search of a market. It is an existing cost that several industries carry under different names, none of which is computed as a property of the sensing geometry.

Who pays	What they call it	What it costs
Agronomy	uncertainty about the field	Re-scouting, prophylactic applications, yield left standing
Carbon markets	MRV cost	25 to 40% of project budget in sampling[11]
Crop insurance	basis risk	The index is measured at the canopy surface and the loss happens in the volume[18]
Grain storage	late detection	The pocket that was never within reach of a probe
Geostatistics	the nugget	Already quantified, conventionally read as noise[12]

Basis risk is the financial name for the unknown fraction. MRV cost is its accounting name. The nugget is its statistical name and the most rigorous of the five, because unlike the others it is already a number in the literature rather than a line in a budget.

The soil carbon MRV platform market was about $85 million in 2025 and is projected near $638 million by 2034, and the stated driver of that growth is demand for lower cost, higher confidence measurement.[11] That is a market whose thesis is the quantity this report makes computable.

9. What did the market just prove about how this can ship?

Any proposal in this area has to survive a sector that has just finished punishing it. Eighteen agricultural technology companies shut down publicly in 2025, across North America, Europe, Asia and Africa, with sensors and internet of things among the hardest hit classes and controlled environment agriculture the worst.[19] Venture funding into the sector fell roughly 70% over three years.

The reported diagnosis is not technological. It is a structural mismatch between what a solution costs and what a farm can pay, and at least one sensing company that closed was described as technically viable.[20] What recovered was low capital expenditure sensing and decision support, with low adoption friction, integrating with equipment already on the farm.

That sets a hard constraint on this work, and we state it as a design rule rather than as background. **A deliverable here cannot be hardware a farmer has to buy.** The unknown fraction satisfies the rule by construction: it is computation over sensors that already exist, and it is a quantity the parties in §8 already pay for. It also has the property that its marginal cost per additional farm is near zero, which matters because 84% of the world's 570 million farms are under two hectares, working 12% of agricultural land and producing about 35% of the world's food,[21] and a per hectare cost model excludes them by construction.

10. A claim of ours that was wrong, and how it was caught

Twice during this work we wrote that machine vision yield estimation assumes a constant visible-to-occluded fruit ratio and never measures it. That is wrong and it is withdrawn.

The ratio is measured. It is measured per orchard by hand counting sample trees, and the published range across real orchards is the 40.85% to 79.83% band used in §6.[13] Our claim survived two drafts because it was convenient: an unmeasured constant is a cleaner target than a measured one.

It was caught by going to find the number in order to replace a placeholder in an instrument, which is the only reason it was caught at all. Nothing in the argument depended on the claim, and the corrected version is stronger: the measurement exists, it shows the factor is not one number, and the published objection to hand calibration is about the impossibility of choosing a representative tree rather than about arithmetic. A quantity that varies by a factor of two across the population it is applied to is a harder problem than one nobody has looked at.

Two smaller corrections belong here. An earlier version of the modality table in the grain bin instrument carried two invented radii, labelled as placeholders; they are replaced in §4 by reaches derived from measured transport coefficients. And an early statement of the reach argument used a round 1.5 m figure for CO_2, giving a 192 times node reduction; the derived value is 179 times, and the earlier figure has been corrected wherever it was published.

11. What does this ask of the people who run it?

The Human Layer argues that the binding constraint in autonomous systems is the human and the pipeline that produces the competence.[3] Agriculture supplies a second instance with a specific addition.

The role has already changed. A farm service technician installs sensors, calibrates controllers, troubleshoots communication networks and updates software, which is a network operations job under an agricultural title. In the United States 36,830 farm equipment mechanics and service technicians were employed in 2023 against a reported shortage of qualified technicians.[22]

The competency this subject adds is **calibration literacy**: judging whether a sensor mesh is still telling the truth after a season in the weather. The need is quantified. Ion selective pH sensors drift by 0.3 to 0.5 units per year, resistive moisture sensors require recalibration every six to twelve months, and salinity above 2 dS/m reduces resistive sensor accuracy by 30 to 50%.[23]

The consequence follows directly from §2 and is worse than it first appears. **A drifted sensor does not add to the unknown fraction. It adds to the wrongly known fraction**, because the system reports its reading with undiminished confidence. A volume covered by drifted sensors scores as resolved and is not.

We reviewed the published curricula of seven precision agriculture certificate programmes and did not locate calibration literacy or measurement uncertainty content in them; they cover positioning, geographic information systems, soil electrical conductivity, unmanned aircraft and equipment.[24]

12. What is this report careful not to claim?

It does not claim that agricultural sensing is bad or that its practitioners are unaware of short range variation. Carbon projects over sample areally precisely because short range variation dominates; the point is that grid sampling can average over that variation and cannot resolve it.

It does not claim that the three instruments are simulators. Each is a first order model with its assumptions stated in its own text: union of non overlapping spheres in the grain bin, a core representative over a circle of the autocorrelation range in the soil, and a continuous Beer–Lambert canopy in an orchard that is not continuous.

It does not claim the unknown fraction is the only thing worth reporting. A field average is a real answer to a real question and is often the question that matters.

It does not claim novelty for the underlying quantities. Diffusion lengths, variogram nuggets and canopy extinction are all standard. The contribution is reading them as one quantity and computing it where it has not been computed.

Finally, it does not claim that nothing of this kind exists anywhere. Six quantities were searched for and not located, and are listed in §13 as absences of retrieval rather than absences of fact.

13. What would sharpen this, and what did we fail to find?

Searched for and not located:

1. An agricultural sensing system, in any sector, reporting unknown volume fraction as an output.
2. Any energy figure per unit of resolved volume for an agricultural sensor network.

3. A published acoustic detection radius in metres for insects in stored grain. Grain is a strong acoustic absorber with attenuation rising as the square root of frequency,[25] so the row exists in our instrument without a number rather than with an invented one.
4. Deployment cost per node for soil microbial fuel cell powered sensors at field scale. The power result is 2026 and the cost was not located.[8]
5. An agricultural data standard carrying measurement uncertainty or drift with the reading. ISOBUS and AgGateway ADAPT move the data; the payload has no uncertainty field.[26]
6. A protocol requirement that a soil carbon project measure the autocorrelation range of the field it is sampling.

Three measurements would sharpen the argument most. A bench measurement of magnetic induction and acoustic reach in soil at agricultural moisture contents would close the fourth barrier the way §4 closed the first. A single field trial reporting resolved and unknown volume alongside its conventional output would establish the reporting convention. And a variogram fitted at sub metre lags on a field already under a carbon contract would turn the nugget from a fitted parameter into a measured one.

14. Conclusions

1. The unknown fraction of a sensed volume is computable from sensing geometry alone, and equals the probability of missing a localized event in that volume.
2. In agriculture the barrier and the signal are the same substance, which is why the quantity is both large and hard to reduce.
3. Three settings with unrelated physics separate the same two questions. The field average is well resolved in all three; the localized answer is not resolved in any of them.
4. Reach, not sensor count, is the variable that moves it, because count and energy both fall with the cube of the radius. In a grain bin this is exact: CO_2 reaches 5.6 times further than heat and needs 179 times fewer nodes, at every time horizon.
5. Four industries already pay for this quantity under four names, and a fifth name for it has been published in every variogram fitted for decades.
6. The 2025 shutdowns make the deliverable form non negotiable: computation over existing sensors, sold to a party that already pays for confidence.

15. The forcing function

Reporting what a system cannot see is not a limitation to be admitted. It is a measurement to be taken, and taking it changes what can be built on top.

The engineering change is specific and available now. Modality reach in the metre range, rather than the centimetre range, collapses the node count and the energy by the cube of the ratio, and in the one case where both transport coefficients are measured the improvement is 179 times. Magnetic induction and acoustic carriers are the candidates in soil, and their reaches are the measurement §13 asks for.

What becomes possible when a sensing system reports its own unknown fraction is a farm that states its own uncertainty. A verifier can price a credit against the volume actually resolved rather than the volume claimed. An insurer can quote basis risk as a computed number rather than a modelled one. A grower can be told where the next sensor buys the most, which is the only sensor question that has ever mattered on a thin margin. And a technician can be trained to know when the mesh has stopped telling the truth, which is a competency, which means it can be taught.

References

1. Institute for Physical AI @ BMI, *OmniSense: The World Model as Projected Perception in a Volume*, Technical Report TR-2026-08. COMPANION REPORT, READ IN FULL
2. R. Gomez, *Physical AI and Logistics Opportunities in Agriculture*, Technical Report TR-2026-32. COMPANION REPORT, READ IN FULL
3. C. R. Glass, *The Human Layer*, Technical Report TR-2026-18. COMPANION REPORT, READ IN FULL
4. Wireless underground sensor network literature on radio attenuation in soil, including magnetic induction and acoustic alternatives. REPORTED, SECONDARY
5. US EPA CLU-IN, *Ground Penetrating Radar*, technology description, penetration depth by material. REPORTED, AGENCY REFERENCE
6. Coupling SAR and optical remote sensing for soil moisture retrieval over dense vegetation, PLOS ONE. REPORTED, PEER REVIEWED
7. Energy efficient LoRa based wireless sensor networks for precision agriculture, MDPI Sensors 23:6332. REPORTED, PEER REVIEWED
8. Soil microbial fuel cell powering buried sensors and backscatter communication, Northwestern University, 2026. REPORTED, LABORATORY RESULT

9. Diffusion and production of carbon dioxide in bulk corn at various temperatures and moisture contents, *Journal of Stored Products Research*. MEASURED, PEER REVIEWED
10. Real time CO_2 monitoring for early detection of grain spoilage and mycotoxin contamination. REPORTED, PEER REVIEWED
11. Soil carbon monitoring, reporting and verification cost structure and market size; sampling share of project budget. REPORTED, INDUSTRY AND SECONDARY SOURCES
12. Semivariogram parameters for soil organic carbon: reported ranges and nugget to sill ratios across sites. MEASURED, PEER REVIEWED, SITE DEPENDENT
13. Percentage of visible apples from a single side across orchards, 40.85% to 79.83%. MEASURED, PEER REVIEWED
14. Correction for occluded fruit in tree fruit load estimation by machine vision with deep learning, *Agronomy* 11:347. MEASURED, PEER REVIEWED
15. Canopy light extinction coefficients by leaf angle distribution, plant physiology reference. REPORTED, TEXTBOOK REFERENCE
16. Light interception by apple orchards as a function of leaf area index, discontinuous canopy asymptote. REPORTED, PEER REVIEWED
17. Optimizing soil moisture sensor placement through spatial variability analysis in orchards. MEASURED, PEER REVIEWED
18. Satellite based data for agricultural index insurance, systematic quantitative literature review, *NHESS* 25:913. REPORTED, PEER REVIEWED REVIEW
19. Agricultural technology shutdowns in 2025 by class and region. REPORTED, TRADE PRESS
20. Analyses attributing 2025 shutdowns to business model rather than technology. REPORTED, TRADE PRESS
21. FAO and Our World in Data on farm size distribution and share of global food production. REPORTED, AGENCY AND SECONDARY
22. Precision agriculture and farm service technician demand, *farmdoc daily*, January 2026. REPORTED, EXTENSION PUBLICATION
23. Soil sensor calibration drift rates and salinity effects on resistive sensor accuracy. REPORTED, SECONDARY AND VENDOR TECHNICAL
24. Published curricula of seven precision agriculture certificate programmes at United States community colleges. REVIEWED, PRIMARY PROGRAMME PAGES
25. Acoustic detection of stored product insects: sound transmission and attenuation in grain, USDA ARS. REPORTED, AGENCY RESEARCH
26. ISOBUS and AgGateway ADAPT interoperability scope. REPORTED, STANDARDS BODY