

DIGITAL TRANSFORMATION ACCELERATOR

Technical Report of the Probe: Low-Cost Portable Soil Fertility Sensors: Myth or Reality?

Authors : Oscar Estrada¹, Gresia Ramos², Miguel Lizarazo³ and Daniel Jimenez⁴

1. Summary

This probe evaluated the performance and accuracy of low-cost portable sensors for measuring Nitrogen (N), Phosphorus (P), and Potassium (K) in soil as a rapid alternative to conventional laboratory analyses. The research comprised two stages: an initial evaluation in Guatemala and a subsequent, more comprehensive validation in Colombia.

In Guatemala, three commercially available NPK sensor models were tested using ten soil samples collected from different regions. Sensor readings were compared with laboratory reference analyses through linear regression models, which yielded low determination coefficients (R^2 ranging from 6×10^{-5} to 0.2458) and revealed a strong dependence of the results on soil moisture content. These findings emphasized the importance of controlling soil moisture and developing calibration procedures prior to field applications.

Building on these results, the Colombia experiment assessed the same sensors using 63 soil samples from various regions, standardized at 20% gravimetric moisture. Polynomial regression models (linear, quadratic, and cubic) were applied to evaluate accuracy using R^2 , adjusted R^2 (R^2_{adj}), the coefficient of variation (CV), and RMSE as performance metrics. Results indicated limited predictive capacity, particularly for Phosphorus and Nitrogen (the latter compared against Organic Matter as a proxy), with adjusted R^2 values generally ranging between 0.052 and 0.479. Potassium (K) showed moderate performance, reaching a maximum adjusted R^2 of 0.539. A complementary experiment examining moisture effects (0–40%) confirmed that NPK and EC readings are highly dependent on soil water content.

Overall, both experiments demonstrated that although these sensors can detect general nutrient trends, their use for precision applications such as fertilizer management is not recommended due to low accuracy and strong sensitivity to soil moisture.

2. Introduction

Efficient soil fertility management is a cornerstone of sustainable agriculture. Traditionally, fertilization decisions have relied on laboratory analyses, which, although accurate, are costly, time-consuming, and offer low temporal resolution, limiting their applicability in precision agriculture.

1,2,3,4: Alliance Bioversity & CIAT researchers

Over the past decade, a new generation of low-cost portable sensors has emerged, many based on electrochemical principles (ion-selective electrodes) or frequency-domain reflectometry, promising instantaneous in-field measurements of key nutrients such as N, P, and K. The adoption of this technology could transform agronomic management by enabling high-frequency monitoring and the generation of high-resolution soil fertility maps to support variable-rate input application (Ameer, Ibrahim, Kulsoom, Ameer, & Sher, 2024).

However, there remains significant uncertainty regarding the reliability, accuracy, and repeatability of these devices, particularly under tropical agricultural soil conditions. Factors such as soil texture, pH, organic matter content, and, critically, soil moisture can cause interferences that affect measurement accuracy.

To contribute to the evaluation of these technologies under tropical conditions, this probe encompassed two complementary experiments conducted in Guatemala and Colombia. The first experiment in Guatemala assessed three commercially available NPK sensor models using soil samples from different regions, while the second experiment in Colombia expanded the evaluation to a larger dataset and included a controlled analysis of soil moisture effects. Together, these experiments provide an integrated assessment of sensor performance and their potential applicability for rapid soil fertility diagnosis in tropical agricultural systems.

The specific objectives of this study were to:

- Evaluate the accuracy of three low-cost sensor configurations by comparing their NPK readings against reference laboratory results.
- Determine whether variations in soil moisture influence the NPK and EC readings of a representative commercial sensor.

3. Materials and Methods

3.1. Evaluated Sensors

Four sensor devices were evaluated:

- Sensor 1 and Sensor 2 (S1 and S2): Two units of the same commercial reference “Soil Speed Tester.” These are generic sensors characterized by the use of individual probes for measuring N, P, and K (Figure 1a). Sensor 2 was only included in the evaluation conducted in Colombia.
- Sensor 3 (S3): A system composed of a generic commercial NPK probe, “Soil Tester 3 in 1”, coupled with a custom data logger based on an Arduino platform, assembled in Guatemala by students from Cunori University (Figure 1b & Figure 1c).
- Sensor 4 (S4): A commercial “Soil Tester 7 in 1” sensor, model 4001-BXSZD, designed to measure N, P, K, and Electrical Conductivity (EC), among other variables, using a single multiparametric probe (Figure 1d).

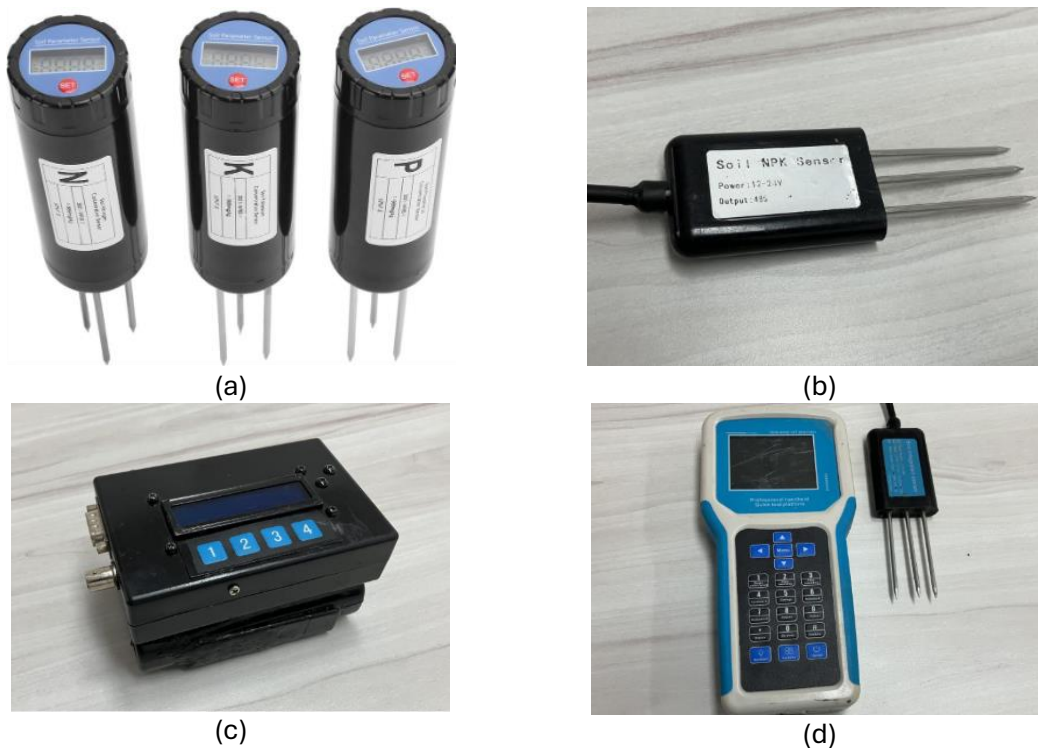


Figure 1. Evaluated sensors. (a) S1 & S2: Soil Speed Tester; (b) S3: Soil Tester 3 in 1 probe; (c) (b) S3: Soil Tester 3 in 1 logger; (d) S4: Soil Tester 7 in 1

Table 1 presents the technical specifications of each sensor, showing that the measurement range and resolution are the same for all three sensors.

Table 1. Technical specifications of the NPK sensors evaluated

Sensor	Cost	Variables	Measurement range	Resolution
S1 & S2 “Soil Speed Tester”	\$USD 96	N,P,K	0-1999 mg/kg	1 mg/kg
S3 “Soil Tester 3 in 1”	\$USD 80	N,P,K	0-1999 mg/kg	1 mg/kg
S4 “Soil Tester 7 in 1”	\$USD 266	N,P,K,pH,EC,Temp,Moisture	0-1999 mg/kg	1 mg/kg

Although there is limited documentation available on the sensors used, it is known that all of them operate based on the Frequency Domain Reflectometry (FDR) principle. This is an electromagnetic technique primarily employed to measure soil moisture (volumetric water content), but some manufacturers incorporate it into multifunctional sensors that also estimate Electrical Conductivity (EC), temperature, and, through empirical correlations, the concentration of N, P, and K ions in solution (Pelletier, Schwartz, Holt, Wanjura, & Green, 2016).

The operating principle is as follows:

- The sensor generates a high-frequency electromagnetic signal (radio frequency) that propagates as a guided wave through the soil (for example, along a metallic rod).
- The signal is reflected depending on the soil's dielectric properties, that is, how the soil responds to the electric field.
- The system measures either the resonance frequency or the reflection time of that electromagnetic wave.
- From this response, the soil's dielectric constant (ϵ) is calculated, which is directly related to volumetric moisture and indirectly to electrical conductivity.

3.2. Soil Samples and Reference Analysis

In the Guatemala experiment, 10 soil samples were collected from five departments representing contrasting agroecological regions: Zacapa, Totonicapán, Quetzaltenango, Huehuetenango, and Chiquimula. The NPK sensors were evaluated by comparing their readings against reference laboratory analyses performed at the Soil Laboratory of CUNORI University.

A total of 63 soil samples were analyzed in Colombia, 50 topsoil samples (0–20 cm) from various agricultural regions of Colombia: 25 from La Jagua de Ibirico (Cesar) and 25 from Tierralta (Córdoba). The Soil Laboratory of the Alliance Bioversity & CIAT supplied the remaining 13 samples. These are standard reference soils routinely used in laboratory calibration and correspond to three different soil types: Llanos Orientales, Santander de Quilichao, and Palmira, with 5, 5, and 3 replicates, respectively.

All Colombia samples had been oven-dried to constant weight and previously analyzed in the organization's soil laboratory to obtain reference values. Phosphorus (P) was determined using the Bray II extraction method. Since the routine laboratory analysis in Colombia did not directly quantify Nitrogen (N) for all samples, the Organic Matter (OM) content was used as a proxy variable for comparison with the N readings obtained from the sensors.

3.3. Experimental Design 1: Accuracy Evaluation

In the Guatemala experiment, the soil samples were analyzed at their original field moisture content, and the results were compared with reference analyses conducted at the soil laboratory. Readings were taken with three sensors (S1, S3, and S4), following the specific protocols of each manufacturer (Figure 2).



Figure 2. Guatemala sampling process

In Colombia, each dry soil sample was rehydrated with distilled water to reach 20% gravimetric moisture. This moisture level was defined based on literature recommendations to ensure adequate ion solubilization without reaching saturation (Sulaeman, Sutanto, Kasno, Sunandar, & Purwaningrahayu, 2024). The samples were covered and stored for 24 hours to achieve homogeneous moisture distribution. Once stabilized, readings were taken with the four sensors (S1, S2, S3, S4), following the specific protocols of each manufacturer (Figure 3). Three readings (replicates) were recorded per sample, and the mean value was calculated for N, P, and K contents.



Figure 3. Colombia sampling process

3.4. Experimental Design 2: Influence of Soil Moisture (Repeated Measures)

This experiment was conducted exclusively in Colombia using the Palmira reference soil and Sensor 4 (NPK + EC). Three replicates (samples A, B, and C) were prepared from this soil. The study followed a repeated-measures design to observe the sensor's response curve under progressive wetting.

Initial readings were taken on all three replicates with oven-dry soil (0% moisture). Distilled water was then added to each replicate to reach 5% gravimetric moisture. The samples were covered and left to rest for one hour to allow stabilization and mineral dissolution. After this resting period, measurements of N, P, K, and EC were taken. Immediately afterward, the process was repeated: additional water was added to reach 10% moisture, the samples were allowed to rest for one hour, and readings were taken again. This wetting–resting–measurement cycle was repeated for all target moisture levels: 5%, 10%, 15%, 20%, 25%, 30%, 35%, and 40% gravimetric moisture.

3.5. Statistical Analysis

For Experiment 1, the relationship between sensor readings (predictor variables) and laboratory results (response variables) was evaluated using regression models. Linear regression was applied for the Guatemala data, while for the Colombia data, polynomial regression models, linear (1st order), quadratic (2nd order), and cubic (3rd order), were used. Model performance was assessed using the Coefficient of Determination (R^2), Adjusted R^2 (R^2_{adj}) (to penalize model complexity), and the Root Mean Square Error (RMSE). Additionally, the Coefficient of Variation (CV) was calculated across three replicate readings per sample to determine the consistency of sensor measurements. Statistical analyses were performed using the scikit-learn and statsmodels libraries in Python.

For Experiment 2, a graphical analysis was conducted to evaluate the sensor response to increasing soil moisture levels.

4. Results and Discussion

4.1. Sensor Performance and Accuracy (Experiment 1)

For Guatemala experiment, the linear regression analysis between sensor readings and laboratory reference values enabled the evaluation of each sensor’s predictive capability in estimating soil nutrient contents (N, P, and K). For each sensor–nutrient combination, the strength of the relationship was assessed through the coefficient of determination (R^2). Table 2 summarizes the sensor readings obtained under different soil moisture levels, highlighting the influence of water content on measurement variability.

Table 2. Variation of N, P, and K with soil moisture percentage in the Guatemala experiment

Soil texture	Soil volumetric moisture (%)	S1 (mg/kg)			S3 (mg/kg)			S4 (mg/kg)		
		N	P	K	N	P	K	N	P	K
Sandy loam	33.95	46	66	157	46	64	129	45	63	126
	26.49	45	45	121	42	60	120	44	62	125
	18.77	12	14	38	28	40	81	33	48	96

Figure 4 presents the linear regression results between sensor readings and laboratory reference values for Nitrogen (N), Phosphorus (P), and Potassium (K). The figure displays a composite mosaic including the nine regression plots corresponding to each sensor–nutrient combination, illustrating the low determination coefficients (R^2) and the limited agreement between sensor and laboratory measurements.

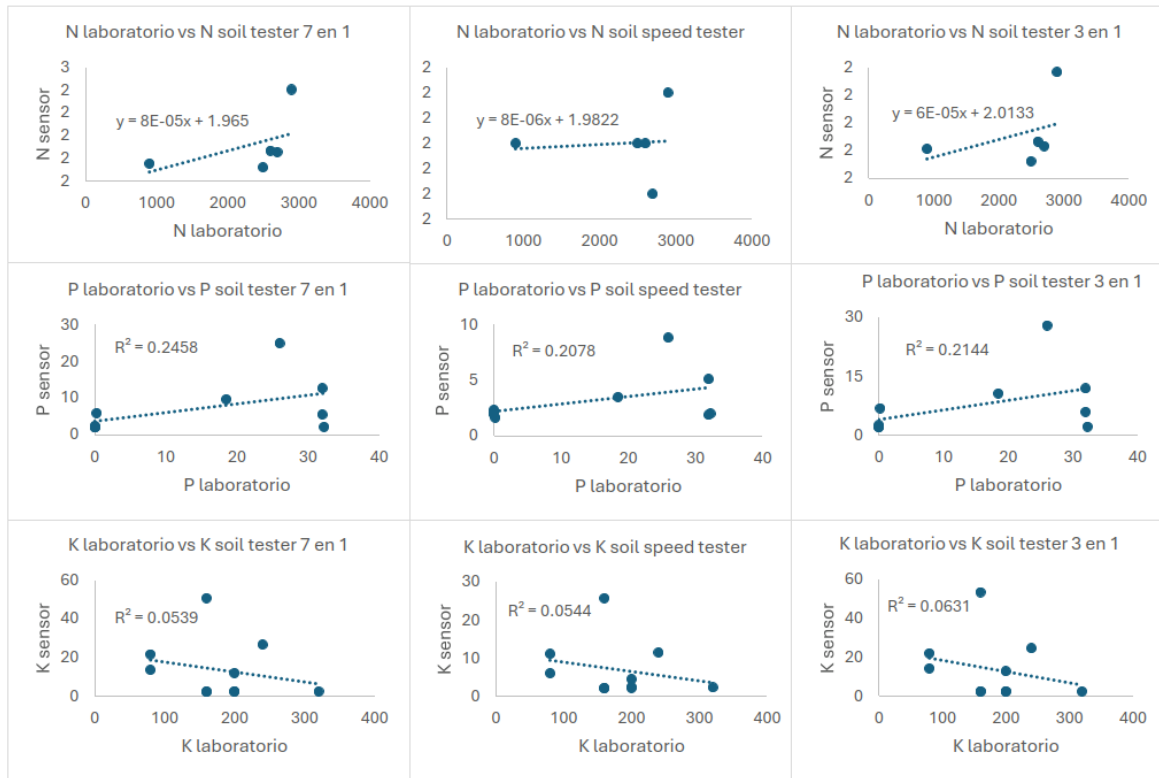


Figure 4. Comparison of N, P, and K readings from sensors S4 (Soil Tester 7 in 1), S1 (Soil Speed Tester), and S3 (Soil Tester 3 in 1) versus laboratory analysis values

Figure 4 shows that the N, P, and K values (in mg/kg) obtained from the three evaluated sensors exhibited determination coefficients (R^2) ranging from 6×10^{-5} to 0.2458, indicating a weak relationship between the sensor readings and the laboratory reference values. Consequently, the statistical models generated were unable to accurately predict the relationship between both variables.

Building upon these initial results, the subsequent experiment conducted in Colombia expanded the evaluation to a larger and more diverse set of soil samples, incorporating controlled moisture conditions. In this stage, an additional unit of Sensor 1 (designated as S2) was included to assess measurement consistency within the same sensor model. The polynomial regression analysis between sensor readings and laboratory reference values allowed evaluation of each sensor's predictive capability in estimating the soil's main nutrient

contents (N, P, and K). For each sensor–nutrient combination, the polynomial model that maximized the adjusted coefficient of determination (R^2_{adj}) was selected, seeking a balance between statistical accuracy and model complexity.

A key finding was that sensor calibration is complex, requiring nonlinear models (2nd or 3rd order) in 12 out of 13 combinations to achieve the best fit (Table 3). With the exception of S4_N, all selected calibration models were statistically significant ($p < 0.05$).

Table 3. Statistical analysis results for the Colombia experiment

Sensor	Var	Lab var	Best degree	R^2_{adj}	RMSE	f p-value	CV %
S1	N	MO	3	0.344	4.42	3.80E-05	10.89
S1	P	P	3	0.195	63.77	0.0013943	11.24
S1	K	K	3	0.539	0.05	3.49E-09	13.38
S2	N	MO	3	0.372	4.39	1.17E-05	12.03
S2	P	P	2	0.181	64.93	0.00102	19.23
S2	K	K	1	0.333	0.05	2.21E-06	16.42
S3	N	MO	3	0.38	6.64	1.36E-06	11.64
S3	P	P	3	0.473	50.4	9.75E-09	13.8
S3	K	K	3	0.406	0.06	4.45E-07	10.71
S4	N	MO	2	0.052	4.38	0.06679652	11.24
S4	P	P	3	0.479	49.01	2.60E-08	11.14
S4	K	K	2	0.43	0.06	5.89E-08	9.64

Potassium (K):

Potassium was the nutrient with the most successful calibration (Figure 5). Sensor S1 achieved the highest calibration accuracy in the entire study ($R^2_{adj} = 0.539$; Figure 5a) using a 3rd-order polynomial model, explaining nearly 54% of the laboratory variance, with moderate repeatability between readings (CV = 13.38%). This result indicates a nonlinear response behavior of the sensor to variations in potassium concentration, justifying the use of higher-order polynomial models.

Sensor S4 showed slightly lower fit quality ($R^2_{adj} = 0.430$; Figure 5d); however, this result was reinforced by greater repeatability, making it the most consistent device in the study (CV = 9.64%). Sensor S3 also demonstrated moderate accuracy ($R^2_{adj} = 0.406$; Figure 5c) with high repeatability (CV = 10.71%). In contrast, Sensor S2 performed the weakest (Figure 5b), showing poor repeatability (CV = 16.42%), which likely contributed to its low calibration accuracy ($R^2_{adj} = 0.333$).

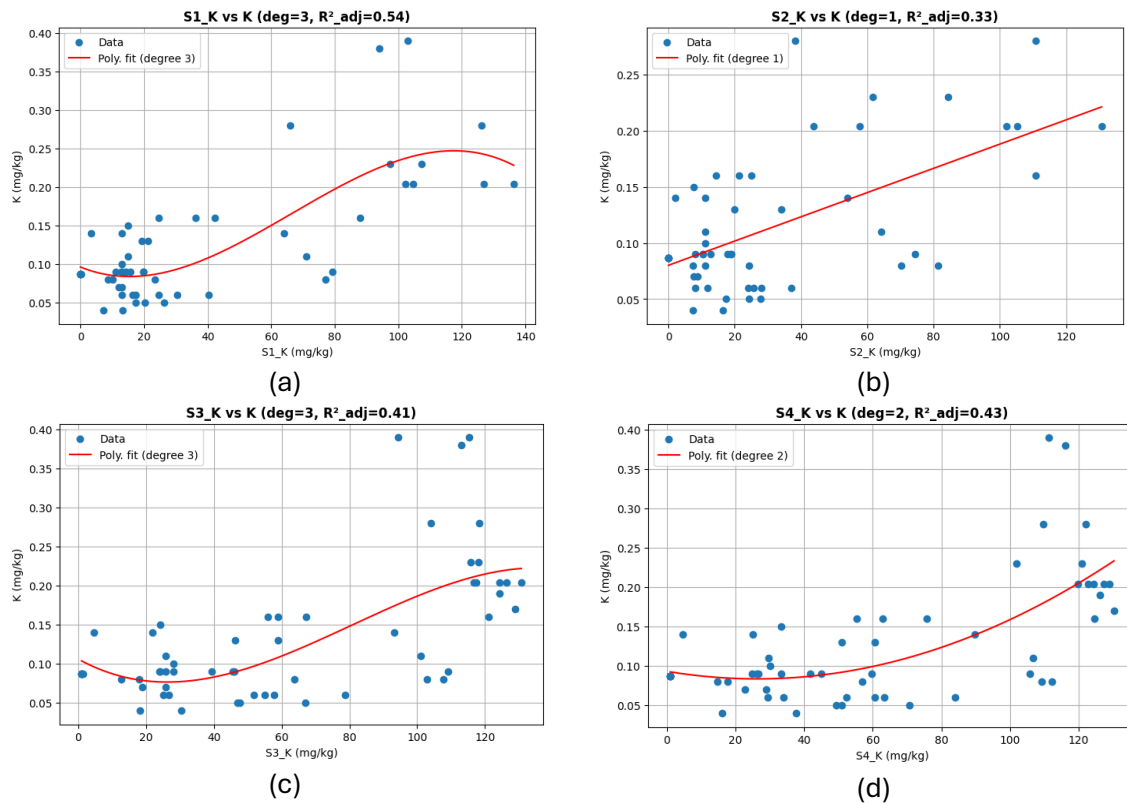


Figure 5. Comparison of potassium (K) measurements among the four sensors. (a) S1; (b) S2; (c) S3; (d) S4

Overall, the results indicate a low to moderate predictive capacity of the sensors across the range of soils evaluated. The finding that Potassium (K) yielded the best calibration metrics is consistent with personal communications from Cenicaña (the national center for sugarcane research, located in Colombia), based on unpublished parallel studies. These studies reported similar trends with other commercial sensors, showing better performance for K estimation under field conditions, but poor accuracy for N and P measurements.

Phosphorus (P):

For Phosphorus (P) (Figure 6), the “all-in-one” sensors (S3 and S4) produced the best results. Sensor S4 achieved the highest calibration accuracy ($R^2_{adj} = 0.479$; Figure 6d) using a 3rd-order polynomial model, supported by good repeatability across samples (CV = 11.14%). Sensor S3 followed very closely, with nearly identical precision ($R^2_{adj} = 0.473$; Figure 6c) and also based on a 3rd-order polynomial model. In contrast, the sensors with individual probes (S1 and S2) performed poorly for this nutrient (Figure 6a and Figure 6b). The repeatability analysis helps explain the weak performance of S2 ($R^2_{adj} = 0.181$), as it exhibited the most inconsistent readings of the entire study (CV = 19.23%).

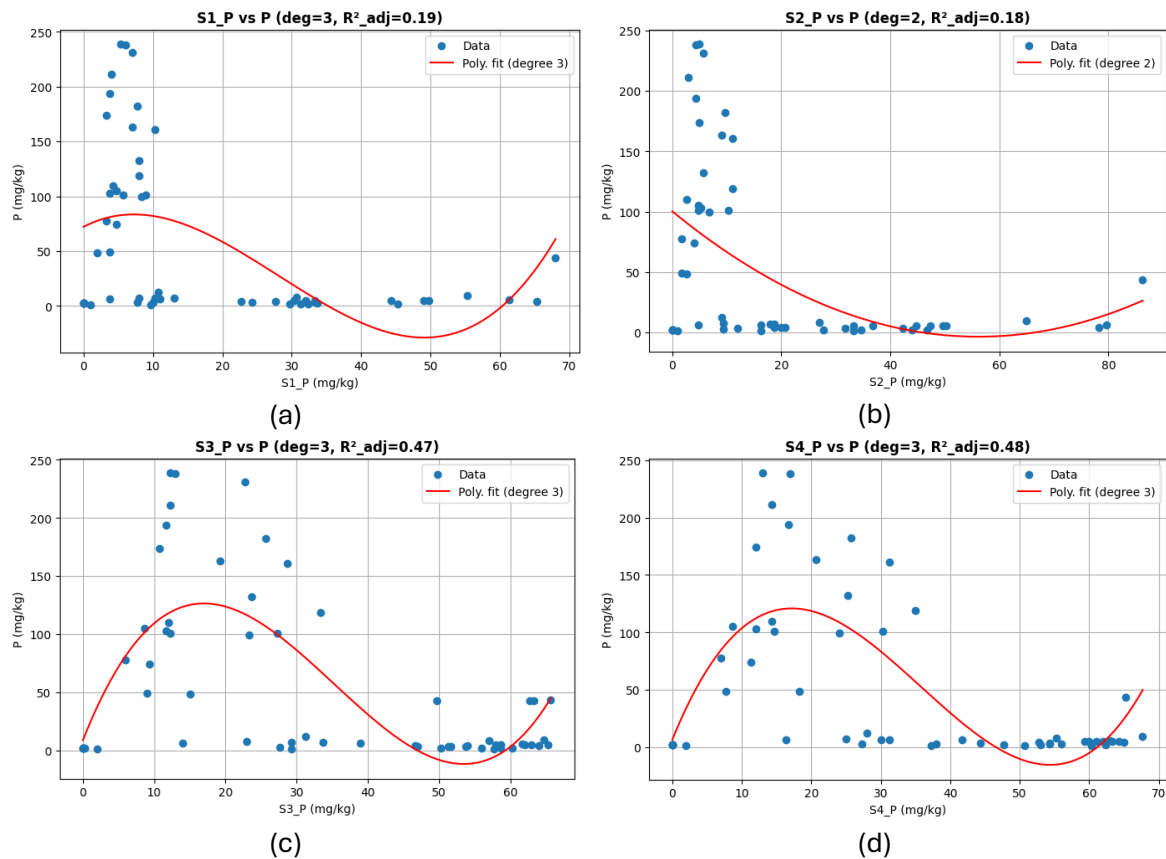


Figure 6. Comparison of phosphorus (P) measurements among the four sensors. (a) S1; (b) S2; (c) S3; (d) S4

The polynomial models corresponding to Phosphorus (P) showed considerably lower adjusted R^2 (R^2_{adj}) values, regardless of the sensor or polynomial degree used. This behavior suggests that the signal captured by the sensors for this nutrient is less stable, likely due to factors such as the low mobility of phosphorus in soil, which introduces additional variability into the measurements. The fitted plots confirm this trend: the polynomial curves exhibit wide dispersion and overlapping among sample readings, limiting their usefulness for precise quantitative estimations.

Nitrogen (N vs. Organic Matter):

Nitrogen, evaluated in correlation with Organic Matter (OM), represented the greatest challenge due to the absence of direct N quantification in the laboratory analyses (Figure 7). Sensor S3 (Figure 7c) and Sensor S2 (individual probe) (Figure 7b) showed “the best” performance, with $R^2_{adj} = 0.380$ (3rd-order polynomial) and $R^2_{adj} = 0.372$ (3rd-order polynomial), respectively. Sensor S4 performed notably poorly in this calibration ($R^2_{adj} = 0.052$, $p = 0.067$; Figure 7d). It is important to emphasize that this was not due to erratic readings, as its repeatability was good ($CV = 11.24\%$); rather, the sensor appears to be consistently inaccurate in predicting N content with OM as proxy.

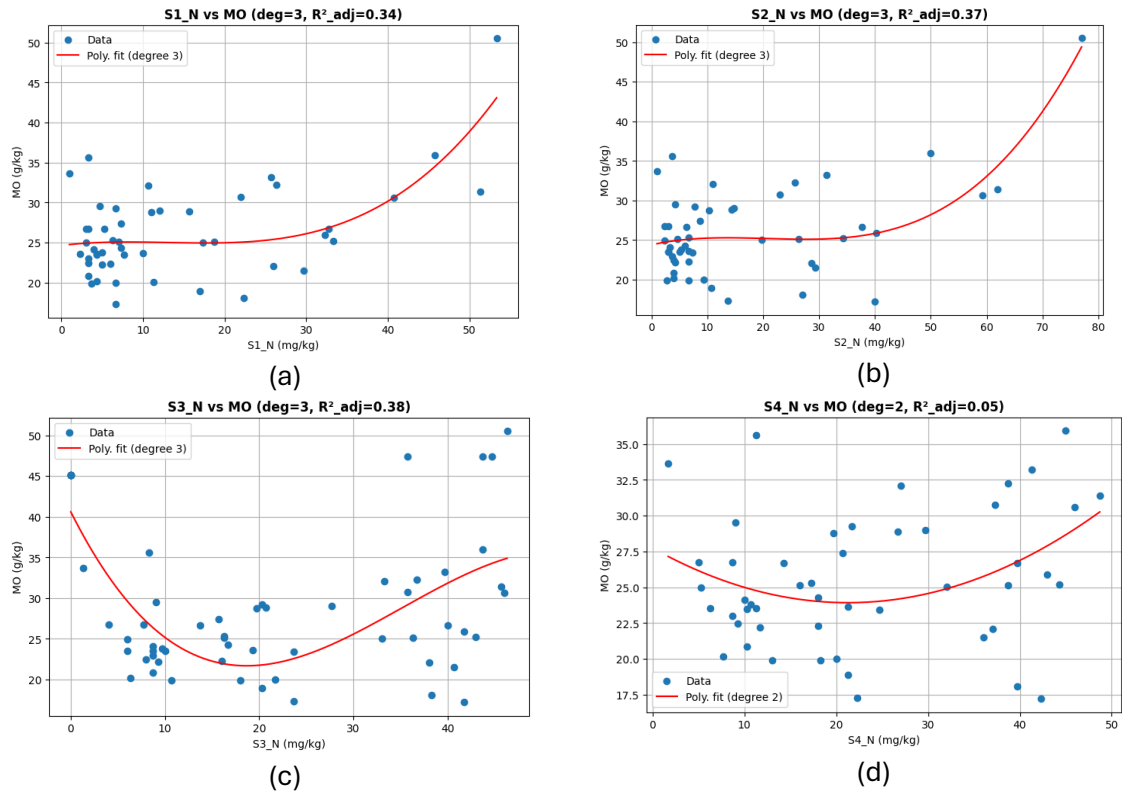


Figure 7. Comparison of nitrogen (N) measurements among the four sensors. (a) S1; (b) S2; (c) S3; (d) S4

Sensor 3 (S3), based on an Arduino system with a single multiparametric probe, showed performance comparable to that of the commercial sensors (S1, S2, and S4) for N and P, but slightly lower for K. This indicates that, although the electronics and signal processing of S3 are simpler, its probe overall response remains within an acceptable range of agreement. Meanwhile, the Sensor 4 (S4), a professional-grade device, did not show substantial improvement in statistical fit compared to the other sensors, suggesting that the main limitations may be associated with the interaction between the electrode and the soil matrix, rather than the precision of the instrument itself.

In all cases, second and third-order polynomial models provided the best compromise between Root Mean Square Error (RMSE) and explanatory capacity, while linear models tended to underestimate variability. However, even with polynomial adjustments, the proportion of explained variance remained below 60%, highlighting the low accuracy of these sensors for quantitative nutrient estimation.

Overall, the results indicate that the evaluated sensors exhibit greater reliability for Potassium (K) and marked limitations for Phosphorus (P) and Nitrogen (N). No significant differences were observed between commercial sensors and the Arduino-based system, reinforcing the potential use of low-cost solutions for exploratory or relative monitoring applications, provided they are supported by adequate local calibration.

4.2. Influence of Soil Moisture (Experiment 2)

Experiment 2 demonstrated that soil moisture content is a critical factor that drastically modulates sensor readings. As shown in Figure 8, the readings of N, P, K, and EC from Sensor 4 increased progressively as gravimetric moisture rose from 0% to approximately 30%.

Beyond 30% moisture, readings for all measured parameters stabilized, showing only marginal or negligible increases when moisture was raised to 35% and 40%. This suggests the existence of a saturation point in the soil solution, where the concentration of measurable ions becomes constant. However, visible soil saturation was observed beginning at around 25% moisture. Readings on oven-dry soil (0% moisture) were zero or near zero for all parameters.

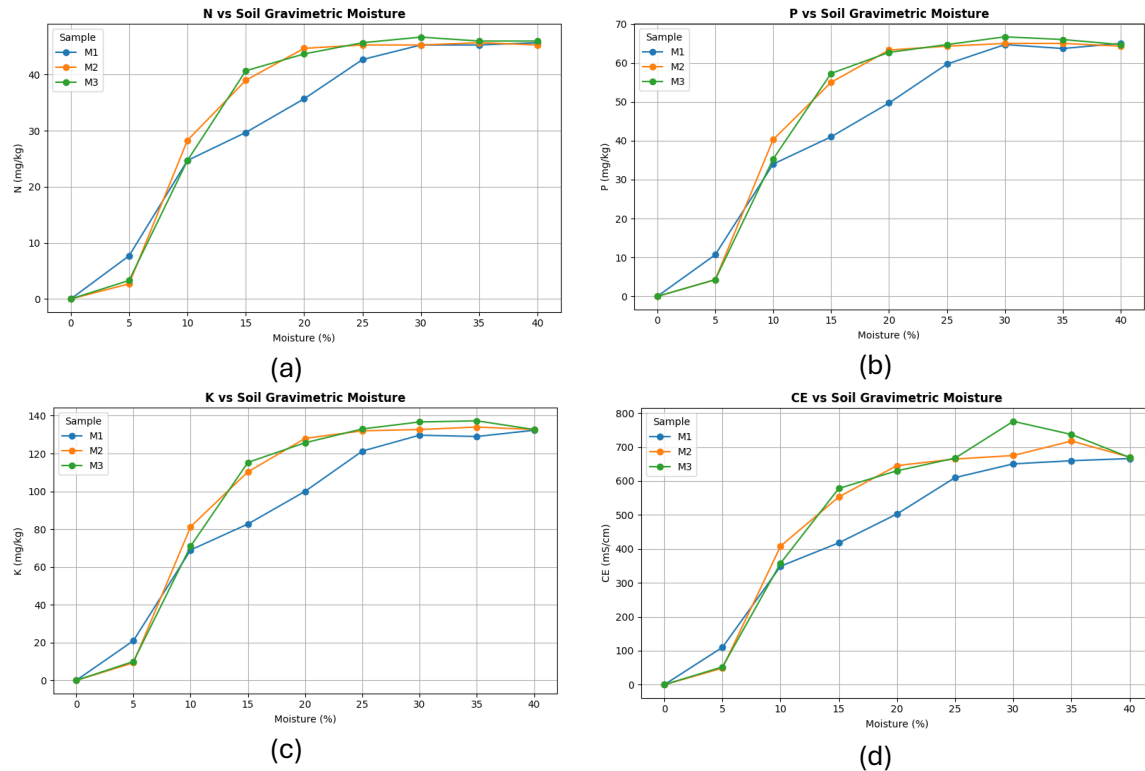


Figure 8. Variation of sensor response with increasing soil moisture

The strong dependence on soil moisture confirms the fundamental operating principle of these sensors: they detect ions dissolved in the soil solution. In the absence of water (0% moisture), there are no free ions available, and therefore, no measurable signal is produced. The progressive increase in sensor readings up to approximately 30% moisture can be explained by the increasing dissolution of soil minerals as the water content rises.

This behavior has a critical practical implication: measurements taken under different field moisture conditions, for example, at different times of the day, cannot be directly compared, as they may lead to misinterpretations of soil fertility status. Standardizing the soil moisture

content, as was done in Experiment 1 (20% gravimetric moisture), is therefore essential to ensure repeatability and comparability among readings.

It is worth noting that this experiment was designed as a repeated-measures analysis, where dissolution time (1 hour per level) was correlated with increasing soil moisture. However, the clear stabilization of readings beyond 30% moisture indicates the occurrence of a true saturation point in the soil solution, rather than an artifact of measurement time.

5. Conclusions and Recommendations

The main conclusion of this probe is that, although the evaluated soil sensors show a statistically significant correlation with laboratory values for certain nutrients, their calibration models do not possess sufficient accuracy for quantitative agronomic decision-making, such as the prescription of fertilizer doses. The robustness of a calibration model is measured by its ability to explain variance (adjusted R^2) and its real-world prediction error (RMSE). In this study, the best calibration scenario in Colombia ($R^2_{\text{adj}} = 0.539$ for S1-K) indicates that the model fails to explain nearly half (46%) of the variability observed in the laboratory data. This large amount of unexplained variance directly translates into an unacceptably high Root Mean Square Error (RMSE). In practice, a high RMSE means that any individual prediction in nutrient units (e.g., mg/kg) carries an excessively wide margin of error. Basing a fertilization recommendation, for example, which entails economic costs and environmental impact, on a value with such a large confidence interval is agronomically unfeasible and risky.

This study also demonstrated that the sensor response is complex, requiring nonlinear polynomial models (2nd and 3rd degree) for optimal fitting, which rules out the use of simple linear calibrations. While Potassium (K) proved to be the most viable nutrient to estimate (R^2_{adj} from 0.40 to 0.54), and Phosphorus (P) showed potential with the “all-in-one” sensors S3 and S4 (R^2_{adj} around 0.47), the estimation of Nitrogen (via Organic Matter) failed completely. The failure of the commercial sensor S4 for Nitrogen ($R^2_{\text{adj}} = 0.052$) underscores that good repeatability among samples (low CV) does not guarantee calibration accuracy.

In the Colombia experiment, it was observed that soil moisture is a critical factor affecting sensor readings. Measurements are zero or near zero when there is no soil moisture and increase nonlinearly with moisture content up to a saturation point (approximately 30% gravimetric). This behavior was also evident in the Guatemala experiment, where variations in moisture levels led to noticeable differences in the N, P, and K values recorded by the sensors. Consequently, field readings taken under different moisture conditions are not comparable and may lead to misinterpretations.

Since the calibration models (low R^2_{adj} and high RMSE) are not robust enough for quantitative prescription, the use of these sensors as a direct replacement for laboratory analysis to determine soil nutrient content is strongly discouraged. Instead of being used for prescription, the main recommendation is to redefine their role toward diagnosing spatial variability (zoning). The fact that the models are statistically significant and that R^2_{adj} values are not zero

(ranging between 0.3 and 0.5) demonstrates that the sensors do capture trends and relative differences in the field. Therefore, their real value lies in their use for mapping fields, rapidly identifying management zones (“high,” “medium,” and “low” areas), and optimizing soil sampling strategies.

6. Acknowledgments

The authors express their sincere gratitude to Dr. Mayesse Da Silva and Eng. Martín Cepeda from the Clima LoCa Project at the Alliance Bioversity & CIAT for providing the soil samples used in the Colombia experiment and for their valuable technical support during the laboratory measurement process.

7. References

- Ameer, S., Ibrahim, H., Kulsoom, F., Ameer, G., & Sher, M. (2024). Real-time detection and measurements of nitrogen, phosphorous & potassium from soil samples: a comprehensive review. *Journal of Soils and Sediments* 24, 2565–2583.
- Pelletier, M. G., Schwartz, R. C., Holt, G. A., Wanjura, J. D., & Green, T. R. (2016). Frequency domain probe design for high frequency sensing of soil moisture. *Agriculture*, 6(4), 60.
- Sulaeman, Y., Sutanto, E., Kasno, A., Sunandar, N., & Purwaningrahayu, R. D. (2024). Developing and Testing a Portable Soil Nutrient Detector in Irrigated and Rainfed Paddy Soils from Java, Indonesia. *Computers*, 13(8), 209.