SJSS SPANISH JOURNAL OF SOIL SCIENCE | SECS SOCIEDAD ESPAÑOLA DE LA CIENCIA DEL SUELO

ORIGINAL RESEARCH
published: 18 September 2025
doi: 10.3389/sjss.2025.14838

# Proficiency Testing for Soil Fertility Analysis

*Asunción Usón Murillo*[1]*, *Jesús Betrán Aso*[2] *and Manuel Sampériz Sarvisé*[1]

[1]*Department of Agricultural and Environmental Science, University of Zaragoza, Zaragoza, Spain,* [2]*Agro-Environmental Laboratory (Government of Aragon), Zaragoza, Spain*

In the context of the FAO Global Soil Partnership, the Spanish Society of Soil Science (Sociedad Española de la Ciencia del Suelo, SECS), the University of Zaragoza and the Agroenvironmental Laboratory of the Government of Aragon promoted initiatives for the harmonization of soil analysis methods in Spain. These included the development of an inventory of laboratories and the organization of proficiency tests of soil analytical results. The first test, carried out in 2019, showed significant discrepancies among laboratories, which led to methodological improvements for the second test in 2021. Twenty-six laboratories participated in the latter test, evaluating soil fertility parameters (organic matter, assimilable phosphorus, potassium and magnesium) and textural fractions (clay and sand). The analyses, carried out in triplicate, were evaluated using robust statistics and Zscore. The 2021 results showed that sample 'HUERTO' had a higher percentage of satisfactory results than "BIPEA," mainly due to its greater standard deviation. The assessment identified problems arising from differences in analytical methods, especially for organic matter. A comparison of the two exercises revealed that, out of 17 laboratories analyzed, 10 improved their overall performance, 3 maintained it and 4 deteriorated. A positive trend was observed for assimilable phosphorus and magnesium, parameters that achieved satisfactory overall ratings, while sand and organic matter content showed more problematic results. Despite the improvements observed, only five out of the seventeen laboratories achieved an overall satisfactory rating, which underlines the need to regularly maintain and strengthen these proficiency tests. These results are particularly relevant within the context of the CAP Strategic Plan, which requires reliable data on the state of agricultural soils. This exercise has proven to be an effective tool for encouraging continuous improvement in the analytical quality of soil laboratories, contributing to the sustainable governance of soil resources.

**Keywords: organic matter, phosphorus, potassium, magnesium, soil laboratories**



***Edited by:***
*Claudia Di Bene, Council for Agricultural Research and Agricultural Economy Analysis | CREA, Italy*

***Correspondence**
*Asunción Usón Murillo, mauson@unizar.es*

***Received:*** *30 April 2025*
***Accepted:*** *29 August 2025*
***Published:*** *18 September 2025*

***Citation:***
*Usón Murillo A, Betrán Aso J and Sampériz Sarvisé M (2025) Proficiency Testing for Soil Fertility Analysis. Span. J. Soil Sci. 15:14838. doi: 10.3389/sjss.2025.14838*

## INTRODUCTION

In December 2012, the Food and Agriculture Organization of the United Nations (FAO) established the Global Soil Partnership (Soil Partnership), with the main objective of improving the governance and sustainable management of soils. This Soil Partnership has five pillars, the fifth of which seeks to harmonize methods, measures and indicators for the sustainable management and protection of soil resources. Within this framework, technical networks were created, including "The Global Soil Laboratory Network" (GLOSOLAN) which was established at an FAO meeting at its headquarters in

Rome, in November 2017 to strengthen the capabilities of soil testing laboratories and respond to the need to harmonize soil testing results.

In Spain, the Spanish Society of Soil Science (Sociedad Española de la Ciencia del Suelo, SECS) together with the University of Zaragoza and the Agroenvironmental Laboratory (of the Government of Aragon), adopted the GLOSOLAN initiative and initiated the development of two fundamental activities:

- An inventory of laboratories performing soil analysis
- A proficiency test to assess the homogeneity of analysis results and facilitate improvements.

Jaime Porta played a crucial role in the development of soil analysis in Spain. As president of the SECS, he championed the Plant Fertility and Nutrition Section. On 10 April 2019 he inaugurated a meeting in Madrid to share the results of the first intercomparison exercise, which was sponsored by the Ministry of Agriculture, Fisheries and Food. Porta began his career as an engineer in the State Agricultural Laboratory Network (LAR), which he referred to as the García Faure project. In 1972, he became the head of the LAR in Galicia, where he established a soil laboratory. He then moved to the LAR of Ebro, where he was tasked with automating soil analysis. To prepare for this, he was sent by Rafael García Faure to study the automated systems at the INRA laboratory in Arras, France, and the Oosterberg cooperative laboratory in the Netherlands. After his time in Zaragoza, he worked at the LAR in Madrid until late 1977, when he left to join the university. Porta's tenure at the Ebro LAR, now the Agroenvironmental Laboratory of the Government of Aragon, had a lasting impact on the automation of soil analysis. He implemented several "Technicon" devices, which are comparable to modern segmented continuous flow systems, for the serial determination of phosphorus and soil organic matter. He also installed two pieces of "Granulostas" equipment (which have no modern equivalent) in a specialized room, enabling the lab to analyze approximately 100 soil textures per week. Porta's most significant contribution was in the field of quality control, as he anticipated the first laboratory quality protocols by two decades. He introduced several key practices, such as working in batches with multiple internal quality controls and establishing sample exchange networks. Porta was also a founding member of the "soils working group" within the Commission of Official Methods of Analysis of Spain, which was responsible for developing and drafting methods for agricultural analysis.

These official and obligatory methods have been used in the agricultural laboratories that depend on the Ministry. They constitute a fundamental tool for harmonizing the results of the analyses and making them interoperable. These official methods began to be published in 1976 (BOE, 1975), with several subsequent editions, the last of which appeared in 1994 (Ministerio de Agricultura, 1994). The official methods remained in that Commission until the end of the 1990s, when, after the transfers to the Autonomous Regions and the incorporation into the European Economic Community, interest in soil analysis declined, and the Commission ceased to act.

Following this new impetus, Order APA/1044/2024 of 23 September 2024 was published on 1 October 2024, designating the Agroenvironmental Laboratory of the Government of Aragon as the national reference laboratory for agricultural soil fertility analysis ("BOE no 237, 2024). In addition, an inventory of soil analysis laboratories is available (SECS, 2018) and two proficiency tests have been carried out to date, which are the focus of this work.

The first test, carried out in 2019, made it possible to fine-tune the methodology (Usón and Betrán, 2020) and revealed the urgent need to standardize procedures and results. The conclusions also led to the proposal of some modifications for subsequent exercises, such as:

- The reduction of the parameters to be analyzed. Since the majority of the intercomparison exercises carried out are limited to one or a few parameters (Bierer et al., 2021; Buczko et al., 2024; Kweon et al., 2015). The new studies focus on soil fertility and eliminate parameters with low response rates (such as carbonates, water retention and micronutrients).
- The introduction of intra-laboratory replicates to also characterize this variability (Guerrero and Bertsch, 2020).

The second test, completed in June 2022, focused on soil texture parameters, with three granulometric fractions, and optionally silt, which was separated into coarse and fine. The test also examined the main soil fertility parameters: organic matter, mineral nitrogen, assimilable phosphorus, potassium and magnesium. The results obtained by the laboratory were requested in three independent replicates (reproducibility conditions) for each parameter.

With these considerations in mind, the objective of this work is twofold. On the one hand, it aims to analyze the degree of homogeneity in soil fertility analysis results from different laboratories in the most recent year, 2021. Second, it seeks to evaluate the analytical consistency among laboratories in Spain with respect to six key soil fertility parameters in the years 2019 and 2021, and to detect trends of improvement or regression in their performance.

## MATERIALS AND METHODS

For both the 2019 and 2021 proficiency tests, two agricultural soil samples were prepared with the standard soil procedure of drying, disaggregation and sieving at 2 mm. The samples were prepared in bulk and homogenized. The results of the first exercise have already been published (Usón and Betrán, 2020) so only the development of the 2021 exercise is described in detail here.

For the second test, the samples were named "BIPEA" and "HUERTO." In the first case it is a sample that was previously provided to the French organizer by that name, and prepared by them; in the second case, the sample was prepared directly by this organization. In both cases, the samples were taken from the superficial horizon (30 cm) of a cultivated soil.

Aliquots of both samples (approximately 500 g) were sent to a total of 29 Spanish laboratories that had shown interest in participating in the exercise. A total of 26 responded; 17 of them had already participated in the first test, and it is with these laboratories that the comparison was made.

**TABLE 1 |** Parameters for which analysis was requested from the participating laboratories, units in which to express the results, and the proposed reference method according to Spanish official soil analysis methods (BOE, 1975).

| Parameters to be analyzed | | | |
|---|---|---|---|
| **Determination** | **Units/method** | **Determination** | **Units/method** |
| Organic matter | g/100 g Walkley-Black | Granulometry | |
| Nitrates (N-NO3) | mg/kg (N) UV spectrophotometry | Clay (<0.002 mm) | g/100 g Discontinuous sedimentation |
| Assimilable phosphorus | mg/kg (P) P Olsen | Fine silt (0.002–0.020 mm) | |
| Assimilable potassium | mg/kg (K) Cation exchange | Coarse silt (0.020–0.050 mm) | |
| Assimilable magnesium | mg/kg (Mg) Cation exchange | Sand (>0.050 mm) | g/100 g Gravimetric analysis |

Following the WEPAL methodology (WEPAL, 2024), at the same time the sample was sent, each laboratory was provided with an Excel sheet for submitting the results. This Excel sheet specified both the units of measurement and the number of significant figures for the results. In order to maintain confidentiality regarding the origin of the results, which is an essential aspect, the code with the number of the sample sent was only known by the laboratory involved and by the two persons responsible for the organization of the exercise. The code for each Laboratory was a number ranging from 11 to 301.

An analytical profile focused on soil fertility and textural fractions was requested (**Table 1**) and the need to perform the analyses in triplicate, if possible in different batches, and to send the results of the 3 replicates was emphasized.

The received results were not modified in any way, and were treated statistically for each parameter.

A robust statistical analysis was used for the evaluation of the results (Laso Sánchez and García-Patrón Peris, 2009a; Laso Sánchez and García-Patrón Peris, 2009b). This analysis consisted of calculating the median and deviation from the median, and eliminating anomalous data through necessary iterations to finally obtain a mean and "consensus" standard deviation with the accepted values. The acceptance interval was calculated symmetrically from the latter values with a 95% confidence level.

The Zscore estimator was developed to assess the distance of each result from the consensus value, in terms of standard deviation:

$$Zscore = \frac{V_{lab} - V_{cons}}{\sigma_{cons}}$$

Where:
$V_{lab}$ = Value of the laboratory.
$V_{cons}$ = Consensus value (average of accepted values).
$\sigma_{cons}$ = Consensus standard deviation (among accepted values).
The interpretation of the Zscore results is as follows:

- $Zscore \leq 2; Satisfactory$
  This interval includes, with a probability of 95%, all correct scores for that property.
- $2 \leq Zscore \leq 3; Questionable$
  There is a 5% probability that scores with this value belong to the population of correct property scores.
- $Zscore \geq 3; Unsatisfactory$
  A score with this value is highly unlikely to belong to the population of correct property scores.

Finally, as a tool for comparison, a "combined Zscore was calculated for both the set of results issued by each laboratory for each parameter and for all the results issued by each laboratory.

$$ZsC = \sqrt{\frac{\sum Zscore^2}{n}}$$

Where:
ZsC = combined Zscore.
n = number of results issued.
Finally, the results of the two tests were compared for the organic matter content, assimilable potassium, assimilable phosphorus, assimilable magnesium and for the sand and clay textural fractions. The procedure was as follows:

- A new code was assigned to each laboratory from 1 to 17; in this case, the codes were the same for both years, and the distribution of the ZsC data for each year was analyzed.
- Subsequently, the distribution of the data was analyzed and, since these did not follow a normal distribution, a logarithmic transformation was performed to normalize them. Then, two analyses of variance were performed:
  ○First, a one-factor ANOVA was analyzed for each laboratory, with the independent variable being the year in which the analyses were carried out.
  ○The second analysis used a one-factor ANOVA for each parameter with the same independent variable: the year in which the analyses were carried out.

# RESULTS AND DISCUSSION

## Results for the 2021 Proficiency Test

**Table 2** summarizes the results obtained in the year 2021, for each variable (organic matter (%), assimilable P, K and Mg (mg/kg) and clay and sand granulometric fractions (%)), and for each sample (BIPEA and HUERTO) it shows the total number of analyses performed (N), taking into account that each laboratory provided up to three replicates, the number of data points used to obtain the consensus value, the consensus value and the standard deviation of the consensus value. For each parameter and sample, the number of results that were unsatisfactory (absolute value of Zscore greater than 3) or questionable (absolute value of Zscore between 2 and 3) and satisfactory (Zscore less than 2) were also recorded.

The percentage of satisfactory results was found to be higher in the HUERTO sample than in the BIPEA sample, mainly due to the higher satisfactory values in the determinations of organic matter, assimilable potassium and clay. It should be taken into account that for these three parameters the standard deviations

**TABLE 2 |** Results of the organic matter content (g/100 g), P, K and assimilable Mg mg/kg along with clay and sand granulometric fractions (g/100 g) for the BIPEA and HUERTO samples in the second exercise (2021): number of responses received (N), consensus value, standard deviation, number of values required to obtain the consensus value (n), number of satisfactory, questionable and unsatisfactory results.

| | Organic matter (g/100 g) | Assimilable P (mg/kg) | Assimilable K (mg/kg) | Assimilable Mg (mg/kg) | Clay (g/100 g) | Sand (g/100 g) |
|---|---|---|---|---|---|---|
| BIPEA sample | | | | | | |
| N | 74 | 76 | 69 | 70 | 54 | 56 |
| n | 37 | 68 | 34 | 60 | 44 | 37 |
| Consensus value | 2.10 | 16.68 | 217.09 | 116.53 | 26.26 | 31.98 |
| Standard deviation | 0.12 | 3.63 | 8.82 | 25.49 | 3.70 | 2.04 |
| No. of Satisfactory results | 34 | 63 | 34 | 57 | 42 | 36 |
| % of Satisfactory results | 45.9% | 82.9% | 49.3% | 81.4% | 77.8% | 64.3% |
| No. of Questionable results | 5 | 5 | 1 | 5 | 2 | 2 |
| No. of Unsatisfactory results | 35 | 8 | 34 | 8 | 10 | 18 |
| HUERTO sample | | | | | | |
| N | 74 | 76 | 69 | 70 | 56 | 56 |
| n | 58 | 67 | 51 | 65 | 50 | 39 |
| Consensus value | 0.69 | 14.81 | 272.38 | 181.82 | 24.89 | 37.77 |
| Standard deviation | 0.17 | 2.81 | 29.15 | 31.43 | 5.05 | 2.74 |
| No. of Satisfactory results | 55 | 63 | 49 | 63 | 47 | 38 |
| % of Satisfactory results | 74.3% | 94.0% | 71.0% | 90.0% | 83.9% | 67.9% |
| No. of Questionable results | 3 | 4 | 5 | 3 | 6 | 4 |
| No. of Unsatisfactory results | 16 | 9 | 15 | 4 | 3 | 14 |

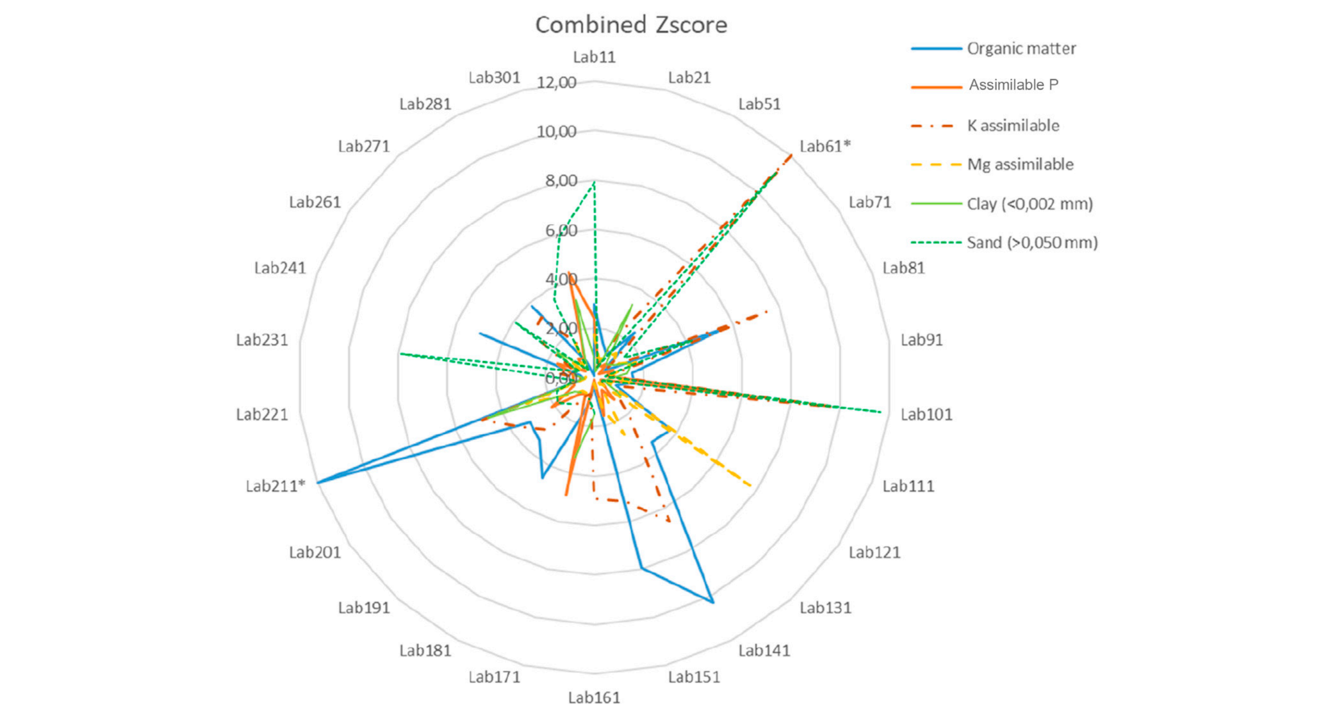


**FIGURE 1 |** Combined Zscore for each laboratory for all analyzed parameters (Organic Matter, Phosphorus, Potassium, Magnesium, Clay and Sand). * Values outside the range.

were higher for the HUERTO sample and this increased the range of acceptable values. In both samples there were a few questionable values, and, consequently, the unsatisfactory values were found to be higher for the BIPEA sample.

Full information on the Zscore values of each laboratory for each parameter is available in the **Supplementary Material**. A summary of the combined Zscore for each parameter is shown in **Figure 1**.

**Table 3** shows the total number of laboratories participating in the exercise for each parameter, and the number of laboratories with satisfactory, questionable, or unsatisfactory ratings for each evaluated parameter after analyzing the combined Zscore for each laboratory.

**TABLE 3 |** Number of laboratories reporting responses for each parameter (organic matter, assimilable P, K and Mg, along with clay and sand particle size fractions) and number of laboratories falling into each classification (satisfactory, questionable or unsatisfactory) according to their combined Zscore for each parameter.

| | Organic matter (g/100 g) | Assimilable P (mg/kg) | Assimilable K (mg/kg) | Assimilable Mg (mg/kg) | Clay (g/100 g) | Sand (g/100 g) |
|---|---|---|---|---|---|---|
| No. of laboratories responding | 26 | 26 | 23 | 25 | 23 | 23 |
| No. of Satisfactory results | 13 | 21 | 13 | 22 | 19 | 14 |
| No. of Questionable results | 2 | 2 | 1 | 2 | 0 | 1 |
| No. of Unsatisfactory results | 11 | 3 | 9 | 1 | 4 | 8 |

In the WEPAL test (2024) 337 laboratories participated with soil samples and the range of unsatisfactory values was 10%, which is lower than the values in our exercise: 28% for the BIPEA sample and 15% for the HUERTO sample. It should be noted that the laboratories participating in the WEPAL exercise aimed to obtain a certification for their results, whereas our exercise had a different objective.

## Organic Matter in the 2021 Proficiency Test

All laboratories provided a response on the organic matter content of the two samples. The organic matter content was higher in the BIPEA sample than in the HUERTO sample and the standard deviation was lower in the first case. This caused the range of acceptable values to be narrower in the BIPEA sample, resulting in many more unsatisfactory results.

The results of the 11 laboratories with an unsatisfactory combined Zscore for organic matter content are discussed in detail:

- Four laboratories (81, 151, 181, and 211) overestimated the organic matter content in all samples (6).
- One laboratory (191) overestimated some of the replicates of each sample (BIPEA and HUERTO).
- Five laboratories (121, 131, 201, 241, and 271) underestimated the organic matter content in the BIPEA sample only.
- One laboratory (141) underestimated the organic matter content in the BIPEA sample and overestimated it in the HUERTO sample, so it appears that there was a transcription error in the data and that the samples were mixed up.

Of the two laboratories with a questionable combined Zscore for organic matter content.

- One (61) underestimated the BIPEA sample.
- One (11) overestimated only one of the BIPEA sample replicates.

It should be noted that the majority of the laboratories measure oxidizable organic matter content using the Walkley and Black method. However, two laboratories (81 and 151) measured it by calcination, and another laboratory measured total organic matter using the DUMAS method (lab 211). The DUMAS procedure determines the total organic carbon released by calcination, so it is normal for it to overestimate (as do comparable calcination methods) the result of organic matter compared to the Walkley and Black method, which is based on the (non-total) oxidation of carbon. These methods were not separated because only three laboratories reported using them, which is insufficient for a separate analysis. Furthermore, the laboratories involved were aware that they overestimated the results for this particular measurement. In other intercomparison exercises, it was evident that comparing organic matter values obtained by different methods was difficult (for reference, the GLOSOLAN exercise should be considered). The Zscore values were excessively high in the present exercise.

The BIPEA exercise organizer analyzed organic carbon data separately depending on the method used. In the June 2024 test (personal communication, July 2024), 27 laboratories participated in the measurement of organic carbon by the oxidation method, and 67% achieved satisfactory results.

## Assimilable Phosphorus in the 2021 Proficiency Test

All laboratories provided a response on the assimilable phosphorus content of the two samples. It should be noted that one of the conclusions of the 2019 exercise (Usón and Betrán, 2020) was a need to improve the results of assimilable phosphorus determination, and this was the reason why three replicates of each sample were requested.

The assimilable phosphorus content of the two samples did not differ greatly and the responses in terms of data dispersion were also similar. This was the parameter for which the highest number of satisfactory results were obtained for the two samples. When we reviewed the combined Zscore, it was also one of the parameters for which more laboratories obtained a satisfactory classification.

The results of the four laboratories with unsatisfactory or questionable combined Zscore for assimilable phosphorus content are discussed in detail:

- Two laboratories (171 and 301) had unsatisfactory overall results due to overestimation of assimilable phosphorus in the two samples. One laboratory (101) overestimated the assimilable phosphorus content only in the HUERTO sample, but the values were so high that the laboratory was classified as having unsatisfactory results overall.
- One laboratory (11) slightly underestimated the assimilable phosphorus in all samples.
- One laboratory (201) slightly overestimated the assimilable phosphorus in all samples.

The latter two laboratories obtained a questionable classification due to bias, to the extent that this prompted a detailed review of the protocol to try to correct it.

In the June 2024 BIPEA test (personal communication, July 2024) the range of acceptable values for assimilable phosphorus by the Olsen method was narrow (0.047–0.075 g/kg) with 77% of results being satisfactory. In our case the ranges of acceptable values were slightly higher as was the percentage of satisfactory results (84%).

### Assimilable Potassium in the 2021 Proficiency Test

All but two of the laboratories provided a response on the assimilable potassium content of the two samples.

The HUERTO sample had a higher assimilable potassium content and a higher standard deviation, resulting in more values being classified as satisfactory. Nevertheless, in both samples the percentages of satisfactory values were lower than those obtained in the 2024 BIPEA test, which had a narrow range of acceptable values (0.17–0.20 g/kg) and obtained 77% acceptable values. The results of the 10 laboratories with an unsatisfactory combined Zscore for assimilable potassium content are discussed in detail as follows:

- Two laboratories (61 and 101) underestimated the assimilable potassium content in all samples.
- Three laboratories (81, 151, and 211) overestimated the assimilable potassium content in all samples.
- Four laboratories (191, 201, 261 and 271) underestimated the assimilable potassium content in the BIPEA sample only. Despite the international recognition of the BIPEA exercise, the determination of potassium in this sample again yielded results similar to those obtained in the assessment of organic matter. This sample showed greater heterogeneity in results, which can only be attributed to less homogeneity in the prepared soil.
- One laboratory (141) overestimated the assimilable potassium content in the BIPEA sample only.

The two laboratories with questionable combined Zscore for assimilable potassium content (131 and 191) underestimated the BIPEA sample.

### Assimilable Magnesium in the 2021 Proficiency Test

All laboratories, except one, submitted a response regarding the assimilable magnesium content of the two samples.

The HUERTO sample had a higher assimilable magnesium content but the two samples had similar standard deviations, which made the ratings of the values similar in the two samples: many more satisfactory values than questionable or unsatisfactory. Both samples had a higher percentage of satisfactory values compared to those obtained in the BIPEA 2024 test (74%).

The results of the three laboratories with unsatisfactory or questionable combined Zscore for assimilable magnesium content are discussed in detail below:

- The only laboratory with an unsatisfactory result (121) overestimated the assimilable magnesium content of all samples.
- One laboratory with a questionable result (141) overestimated the assimilable magnesium content in one sample only.
- The other laboratory with questionable results (211) slightly overestimated the assimilable magnesium content in both samples.

### Granulometric Fractions in the 2021 Proficiency Test

The soil particle size classes did not exactly match those of other exercises, so in this one we requested four textural classes: sand (2 mm<ø<0.05 mm), coarse silt (0.05<ø<0.02 mm), fine silt (0.02 mm<ø<0,002 mm), and clay (ø<0,002 mm) in the BIPEA exercise, five textural classes were requested; In addition, some laboratories did not send the silt fraction or they sent it unaggregated; therefore, we only analyzed the clay and sand fractions.

All but three of the laboratories submitted a response on the clay and sand content of the two samples.

#### *Clay*

The clay content was slightly higher in the BIPEA sample; however, the HUERTO sample had a higher standard deviation, leading to more unsatisfactory results in the BIPEA sample. In both samples the percentage of satisfactory results exceeded that of the 2024 BIPEA test (72%) although in this case the range of acceptable values was narrower (17.7–21.7 g/100 g) and had greater agronomic application.

The results of the four laboratories with an unsatisfactory combined Zscore for clay content are discussed in detail below:

- Two laboratories (171 and 211) underestimated the clay content in the two samples.
- One laboratory (51) overestimated the clay content in the two samples.
- One laboratory (301) overestimated the clay content in the BIPEA sample only.

#### *Sand*

The sand content was slightly higher in the HUERTO sample, and the standard deviations were small, causing a large number of unsatisfactory results in both samples. In both samples, the percentage of satisfactory results was lower than that of the BIPEA 2024 test (78%) although the range of acceptable values was similar (11.8–14.2 g/100 g).

The results of the eight laboratories with an unsatisfactory combined Zscore for sand content are discussed in detail below:

- Four laboratories (101, 231, 281, and 301) overestimated the sand content in the two samples.
- Two laboratories (81 and 261) underestimated the sand content in the two samples.
- One laboratory (11) overestimated the sand content in the BIPEA sample.
- One laboratory (61) underestimated the sand content in the BIPEA sample.

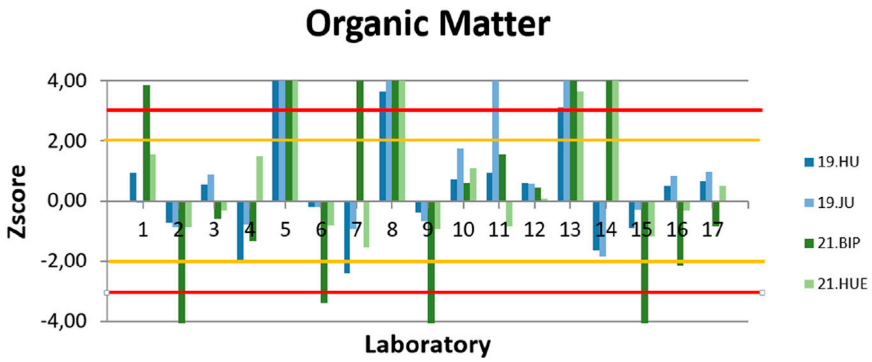


**FIGURE 2 |** Zscore for the results issued by each laboratories in the 2019 and 2021 proficiency tests for the organic matter parameter. In some cases, the Zscore value was found to be higher/lower than the plotted limit for laboratories 2, 5, 8, 10, 13, 14 and 15 for the 4 samples analyzed (HUERTO and JUNTAS in 2019 and BIPEA and HUERTO in 2021). The yellow line represents the lower limit of questionable results, and the red line represents the lower limit of unsatisfactory results.

The laboratory with a questionable rating (141) slightly overestimated the BIPEA sample and slightly underestimated the HUERTO sample.

## Comparison of Results Between the 2019 and 2021 Proficiency Tests

### Organic Matter

All of the laboratories that participated in the two tests (17) delivered organic matter results.

**Figure 2** shows generally higher Zscore values in the second test, especially for the BIPEA sample. Six laboratories obtained satisfactory results in both samples of the two tests. Laboratories 5, 8, and 13 showed unsatisfactory values in the two exercises, and it is possible that these inconsistencies are due to the analytical method used, as other authors (Guerrero and Bertsch, 2020) have referenced these differences, and in other tests, such as WEPAL, the results were analyzed according to the method used. The case of laboratories 2, 14, and 15 was more relevant, as they went from satisfactory results in the 2019 test to unsatisfactory results in the 2021 test. The evolution has not been, in general terms, positive for this parameter.

### Assimilable Phosphorus

In the comparison of Zscore values for phosphorus, all but one laboratory had acceptable average results in the 2021 exercise. Laboratory 11 significantly overestimated the phosphorus concentration (Zscore of 5.00 and 4.73 for each sample) as it did in one of the samples in the 2019 exercise. The rest of the laboratories either maintained or improved their rating in the results for this parameter (**Figure 3**). Globally, a trend toward improved results was observed, in agreement with (Becker et al., 2019) who also found that the repetition of proficiency tests contributed to the improvement of the results.

### Assimilable Potassium

In this case, an apparent worsening of the score rating was evident, with 8 laboratories deemed to be unsatisfactory and 3 deemed to be questionable in 2019, to 13 rated unsatisfactory in 2021 (**Figure 4**). However, it should be noted that for one of the samples of the second test (BIPEA), the range of satisfactory values was very narrow and resulted in high Zscore values with relatively small differences accounting for 8 of the 13 unsatisfactory results.

### Assimilable Magnesium

There was an apparent improvement in the rating of results for assimilable magnesium, going from 6 unsatisfactory and three questionable laboratories in 2019, to one unsatisfactory and two questionable laboratories in the 2021 test (**Figure 5**).

### Granulometric Fractions

The interpretation of the Zscore values in the analysis of the grain size fractions was difficult in both tests. For clay content, the range of acceptable values was very wide in the two tests, leading the majority of laboratories to report acceptable values, with notable exceptions (**Figure 6**). It should be noted that the clay fraction is a much more important determinant of soil behavior and plant response than the other textural fractions (Porta et al., 2003), so values with very different responses were accepted as valid.

Conversely, the sand fraction showed a very narrow range of acceptable values, leading the majority of laboratories to report unsatisfactory results (**Figure 7**).

In addition, fewer results were received for the granulometric fractions in the second test. Specifically, in the case of the clay fraction (**Figure 6**) two laboratories did not send results for this parameter in 2021. For the sand fraction, there were also three laboratories that did not send results in the second test.

### Overall Evolution of Laboratories

**Figure 8** shows the combined Zs-core values (root mean square) of all laboratories for each proficiency test (1 and 2), which helps to interpret their comparison. The ratings went from 7 overall unsatisfactory and 4 questionable laboratories in 2019, to 5 unsatisfactory and 7 questionable in 2021. Although the results are not ideal, 10 of the 17 laboratories analyzed improved their overall Zscore, 3 maintained it, and only 4 clearly worsened it.

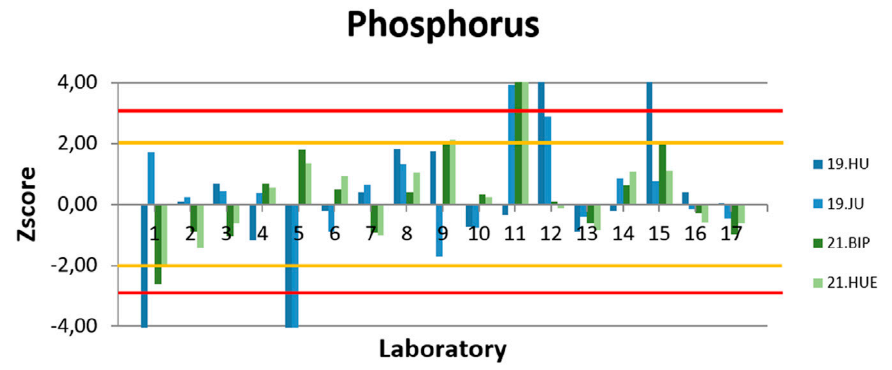


**FIGURE 3 |** Zscore for the results issued by each laboratories in the 2019 and 2021 proficiency tests for the phosphorus parameter. In some cases, the Zscore value was found to be higher/lower than the plotted limit for laboratories 5, 1, 12 and 15 for the 4 samples analyzed (HUERTO and JUNTAS in 2019 and BIPEA and HUERTO in 2021). The yellow line represents the lower limit of questionable results, and the red line represents the lower limit of unsatisfactory results.

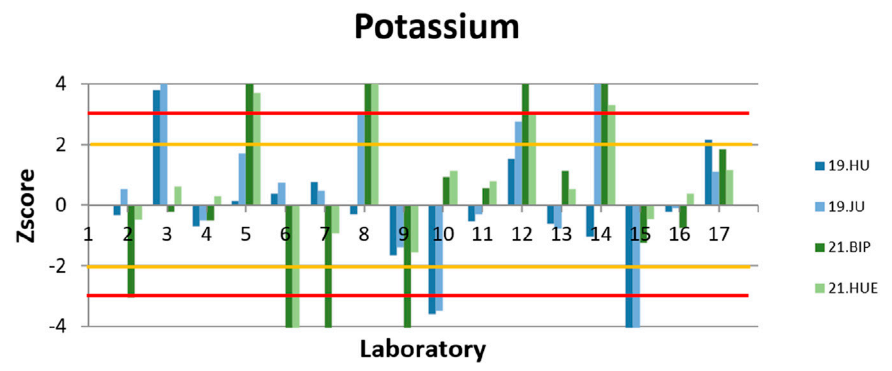


**FIGURE 4 |** Zscore for the results issued by each laboratories in the 2019 and 2021 proficiency tests for the potassium parameter. In some cases, the Zscore value was found to be higher/lower than the plotted limit for laboratories 5, 6, 7, 8, 8, 9, 12, 14 and 15 for the 4 samples analyzed (HUERTO and JUNTAS in 2019 and BIPEA and HUERTO in 2021). The yellow line represents the lower limit of questionable results, and the red one represents the lower limit of unsatisfactory results.

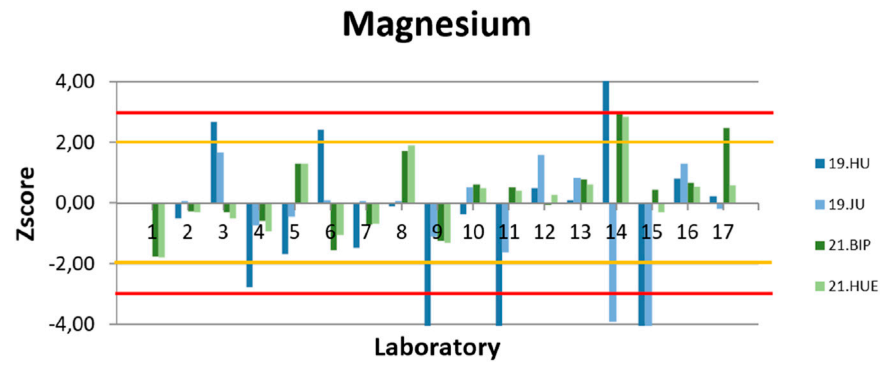


**FIGURE 5 |** Zscore for the results issued by each laboratories in the 2019 and 2021 proficiency tests for the magnesium parameter. In some cases the Zscore value was found to be higher/lower than the plotted limit for laboratories 9.11, 14 and 15 for the 4 samples analyzed (HUERTO and JUNTAS in 2019 and BIPEA and HUERTO in 2021). The yellow line represents the lower limit of questionable results, and the red line represents the lower limit of unsatisfactory results.

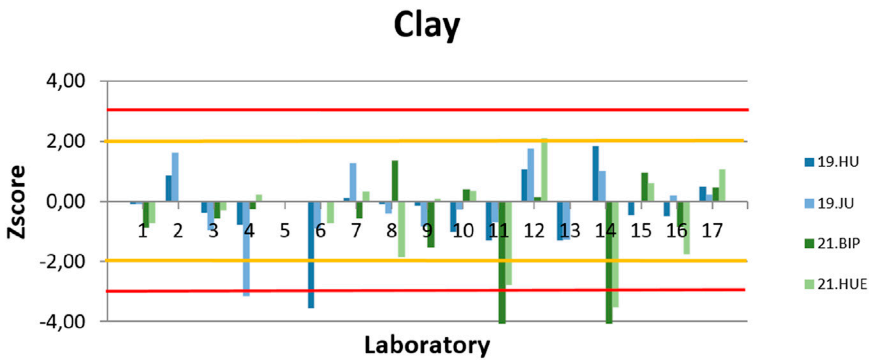


**FIGURE 6 |** Zscore for the results issued by each laboratories in the 2019 and 2021 proficiency tests for the clay parameter. In some cases, the Zscore value was found to be higher/lower than the plotted limit for laboratories 11 and 14 for the 4 samples analyzed (HUERTO and JUNTAS in 2019 and BIPEA and HUERTO in 2021). The yellow line represents the lower limit of questionable results, and the red one represents the lower limit of unsatisfactory results.

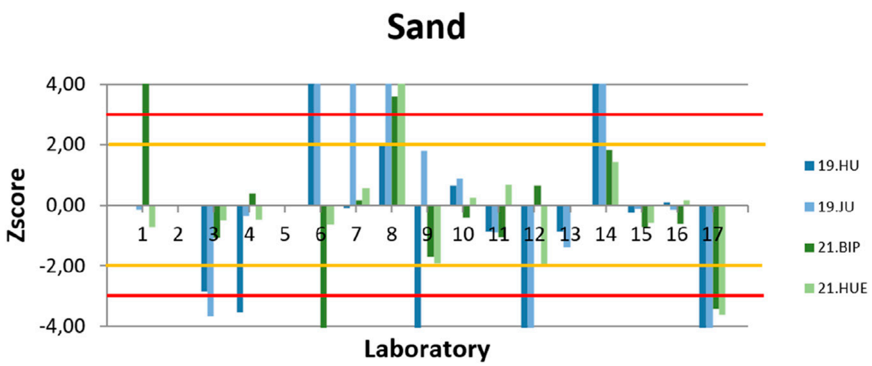


**FIGURE 7 |** Zscore for the results issued by each laboratories in the 2019 and 2021 proficiency tests for the sand parameter. In some cases, the Zscore value was found to be higher/lower than the plotted limit, for laboratories 1, 6, 7, 8, 9, 12, 14 and 17 for the 4 samples analyzed (HUERTO and JUNTAS in 2019 and BIPEA and HUERTO in 2021). The yellow line represents the lower limit of questionable results, and the red one represents the lower limit of unsatisfactory results.

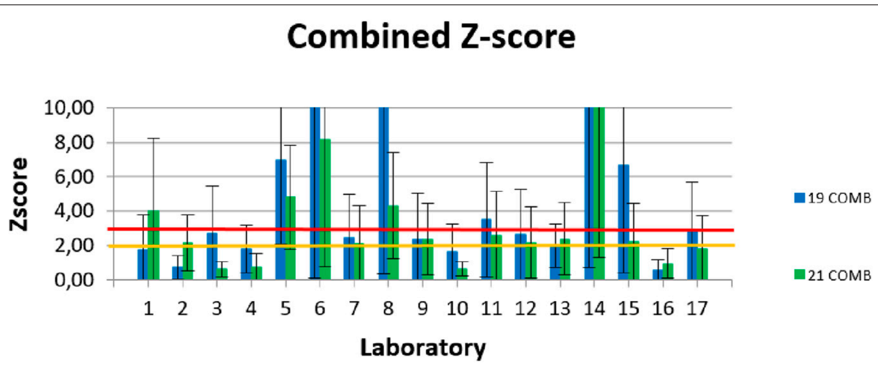


**FIGURE 8 |** Combined Zscore value for the 6 determinations performed (organic matter, phosphorus, assimilable potassium and magnesium and the textural fractions of clay and sand) for each laboratories in the 2019 and 2021 proficiency tests. The attached line at the top represents the variability of the Zscore obtained by each laboratory in each edition of the tests. The yellow line represents the lower limit of questionable results, and the red one represents the lower limit of unsatisfactory results.

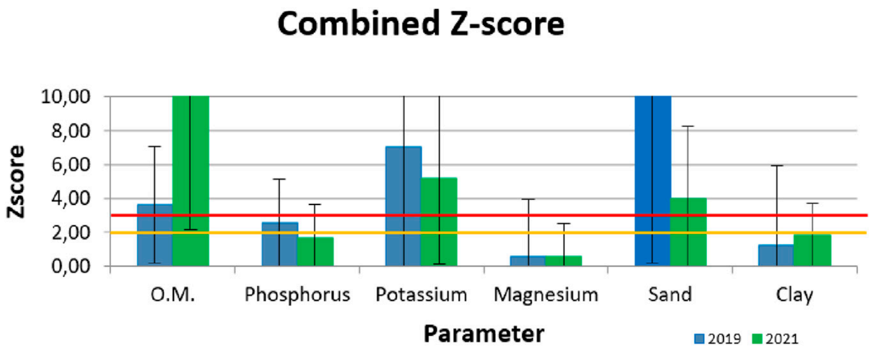


**FIGURE 9 |** Combined Zscore value for the 17 laboratories for the 6 determinations performed (organic matter, phosphorus, assimilable potassium and magnesium and the textural fractions of clay and sand) in the 2019 and 2021 proficiency tests. The attached line at the top represents the variability of the Zscore obtained from each laboratories in each edition of the tests. The yellow line represents the lower limit of questionable results, and the red one represents the lower limit of unsatisfactory results.

In the analysis of variance performed individually for each laboratory, there were statistically significant differences ($P < 0.05$) observed for only two laboratories: laboratories 3 and 4; in both cases, the combined Zscore was lower in the 2021 test compared to the 2019 test.

### Overall Evolution of the Analyzed Parameters

Regarding the evolution of the scores for each parameter, the situation was also observed to be uneven when analyzing the set of results obtained (combined Zscore) (**Figure 9**). Several studies observed that the repetition of these proficiency tests improves results (Becker et al., 2019; Houba et al., 1996) and that they are therefore a good tool for improving the quality of analytical laboratories.

The combined Zscore value in the two tests remained satisfactory for clay and assimilable phosphorus content, which is a considerable achievement, since other studies have reported difficulties in homogenizing results for this parameter (Hanson et al., 1998).

In the two tests, the Zscore values remained unsatisfactory for assimilable potassium content, and sand content went from unsatisfactory results in the 2019 test to questionable results in the 2021 test. It should be noted that the second test used an average of 6 values (instead of 2 in the first exercise) and that the ranges of satisfactory values have decreased.

For the organic matter content, the results were found to be the opposite, going from questionable values in the 2019 test to unsatisfactory in the 2021 test. It should be taken into account that evaluating the organic matter content in Spanish soils is one of the objectives of the environmental monitoring program of the CAP Strategic Plan (MAPAMA, 2023) so the results obtained in this work are relevant for analyzing the 128,000 samples collected for this purpose.

Finally, the estimation of the assimilable magnesium content went from questionable mean values in the 2019 test to satisfactory values in the 2021 test. When performing the ANOVA, this was the only parameter showing significant differences with a $P < 0.10$ ($P = 0.088$). Therefore, it can be stated that the assimilable magnesium determination significantly decreased its Zscore value in the second proficiency test compared to the first one performed.

## CONCLUSIONS

We analyzed the combined Zscore obtained by each participating laboratory on the results of six essential soil fertility parameters. In the 2021 exercise, only five of the 17 laboratories obtained a satisfactory rating, seven received a questionable rating, and another five were found to be unsatisfactory. In 2019, these figures were 6, 4, and 7, respectively. Between 2019 and 2021, 4 laboratories improved their Zscore rating, 4 worsened, and 9 remained the same; of these, 4 remained unsatisfactory and 2 remained questionable.

The variability found among laboratories makes the results obtained a matter of chance, and the interpretation or application of these results can lead to radically wrong decisions. This is unacceptable in the national and European contexts of interest in soil and its conservation.

Analyzing the combined Zscore for the parameters revealed that all laboratories maintained their rating except for the assimilable phosphorus result, which indicated an improvement from questionable in 2019 to satisfactory in 2021. The results for organic matter, assimilable potassium, and sand remained unsatisfactory, while those for magnesium and clay remained satisfactory.

In the second test (2021), the collection of replicates did not provide much information but substantially complicated the management of the exercise. The results obtained showed the need to periodically carry out this type of proficiency test.

Finally, the two intercomparison exercises studied in this paper demonstrated the urgent need to standardize results between laboratories offering soil fertility analysis in Spain. This has been communicated to the Ministry of Agriculture, Fisheries and Food, and the main proposal for the future is the designation of a national reference laboratory in this field. This was already achieved with the designation of the Agro-Environmental Laboratory as the national reference laboratory on 1 October 2024.

Within this framework, the following additional measures are proposed:

- Annual intercomparison exercise. The two-sample format has proven to be convenient and efficient. It should include at least, the usual parameters.
- Sharing of analysis methods, with a view to standardizing their application. This applies at least to those methods that are considered official in Spain that have become obsolete, not so much in terms of extraction techniques as in terms of determination techniques. In practice, this would involve reviving the Official Commission on Analysis Methods, which ceased operations in the early 1990s.
- Attempt to carry out a comparative study of methods. To accomplish this, more information must be required when sending results, and more laboratories with different methods must participate, with some even participating with more than one method.
- Encouraging participation in parameters offered by fewer laboratories, such as saturated paste, microelements, heavy metals, water retention; considering including parameters that are not strictly considered soil fertility parameters or that are rare in the market, such as phytosanitary residues, biological parameters, aflatoxins, etc.

All information obtained should be used to inform the design of future exercises.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

AU funding acquisition to conduct the study, conceived and designed research, collected data, statistical analyses and wrote the manuscript. JB also participated in the design of the research, collected data, performed the figures and tables and wrote the manuscript. MS participated in revising the manuscript and references. All authors contributed to the article and approved the submitted version.

## FUNDING

The author(s) declare that no financial support was received for the research and/or publication of this article.

## CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## GENERATIVE AI STATEMENT

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## ACKNOWLEDGMENTS

The authors would like to express their gratitude to the laboratory staff who participated in the exercises.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontierspartnerships.org/articles/10.3389/sjss.2025.14838/full#supplementary-material

## REFERENCES

Becker, R., Sauer, A., and Bremser, W. (2019). Fifteen Years of Proficiency Testing of Total Petrol Hydrocarbon Determination in Soil: A Story of Success. *Accreditation Qual. Assur.* 24, 289–296. doi:10.1007/S00769-019-01383-X

Bierer, A. M., Leytem, A. B., Rogers, C. W., and Dungan, R. S. (2021). Evaluation of a Microplate Spectrophotometer for Soil Organic Carbon Determination in south-central Idaho. *Soil Sci. Soc. Am. J.* 85, 438–451. doi:10.1002/SAJ2.20165

BOE (1975). *BOE-A-1976-6778 Orden De 5 De Diciembre De 1975 Por La Que Se Aprueban Como Oficiales Los Métodos De Análisis De Suelos Y Aguas*, 6458–6491.

BOE (2024). *BOE-A-2024-19791 Orden APA/1044/2024, De 23 De Septiembre, Por La Que Se Designa Laboratorio Nacional De Referencia En Materia De Análisis De Fertilidad De Suelos Agrarios*, 120669–120670.

Buczko, U., Totubaeva, N., and Kuchenbuch, R. O. (2024). Comparison of the Machigin and CAL Methods for Extraction of Plant Available P in Soils. *Commun. Soil Sci. Plant Anal.* 55, 2217–2231. doi:10.1080/00103624.2024.2345149

Guerrero, A., and Bertsch, F. (2020). *Primer Informe Del Ejercicio De Intercomparación De La Red Latinoamericana De Laboratorios De Suelos-LATSOLAN*. Roma. doi:10.4060/ca9251es

Hanson, D., Kotuby-Amacher, J., and Miller, R. O. (1998). Soil Analysis: Western States Proficiency Testing Program for 1996. *Fresenius J. Anal. Chem.* 360, 348–350. doi:10.1007/s002160050707

Houba, V. J. G., Uittenbogaard, J., and Pellen, P. (1996). Wageningen Evaluating Programmes for Analytical Laboratories (WEPAL), Organization and Purpose. *Commun. Soil Sci. Plant Anal.* 27, 421–431. doi:10.1080/00103629609369565

Kweon, G., Lund, E. D., Maxton, C., Lee, W. S., and Mengel, D. B. (2015). Comparison of Soil Phosphorus Measurements. *Trans. ASABE* 58, 405–414. doi:10.13031/trans.58.10903

Laso Sánchez, J. M., and García-Patrón Peris, A. (2009a). *Evaluación De Resultados De Ensayos De Aptitud*. Madrid: V Iberlab. 1–5.

Laso Sánchez, J. M., and García-Patrón Peris, A. (2009b). *Tratamientos Estadísticos De Aptitud: Aplicación De La Mediana Para Detección De Resultados Anómalos*. Madrid: V Iberlab. 6–11.

MAPAMA. (2023). Programa De Vigilancia Ambiental Del Plan Estratégico De La Pac Para España 2023-2027: Planteamiento General Y Primer Avance De Indicadores De Contexto/Estado Medioambientales Y Climáticos.

Martínez Antonio, M. E. (2021). *Del Laboratorio Agrario Regional Del Ebro Al Laboratorio Agroambiental. 50 Años De Historia*, 1971–2021.

Ministerio de Agricultura (1994). "Métodos Oficiales De Análisis," in *Tomo III*, 662.

Porta, J., López-Acevedo, M., and Roquero, C. (2003). "Edafologia Para La Agricultura Y El Medio Ambiente, 3ª," in *Edafología: Para La Agricultura Y El Medio Ambiente*. Madrid: Mundi-Prensa.

SECS. (2018). Inventario Laboratorios De Suelos. Available online at: https://www.secs.com.es/inventario-%20laboratorios-de-suelos/ (Accessed April 27, 2025).

Usón, A., and Betrán, J. (2020). Puesta a Punto De La Metodología De Ejercicios De Intercomparación Para Análisis De Fertilidad De Suelos. Primeros Resultados. *Span. J. Soil Sci.* 10, 191–197. doi:10.3232/SJSS.2020.V10.N3.01

WEPAL (2024). "International Soil-Analytical Exchange," in *Final Annual Report 2023*, 372.