

Puesta a punto de la metodología de ejercicios de intercomparación para análisis de fertilidad de suelos. Primeros resultados

AUTORES

Usón Murillo A. @
mauson@unizar.es

Betrán Aso J.

@ Corresponding Author

¹Escuela Politécnica Superior-Universidad de Zaragoza. Carretera Cuarte s/n. 22071 Huesca, Spain.

Tuning of the proficiency testing methodology for soil fertility analysis. Preliminary results
Ajuste da metodologia de testes de proficiência para análise de fertilidade do solo. Resultados preliminares

Received: 10.01.2020 | Revised: 24.06.2020 | Accepted: 27.09.2020

RESUMEN

Este trabajo recoge la metodología de realización de un ejercicio de intercomparación entre laboratorios españoles que ofertan análisis de la fertilidad de los suelos. Se prepararon alícuotas de cada una de las dos muestras de gran volumen, tomadas en dos suelos con marcadas diferencias entre sí. Una vez testada la homogeneidad de alícuotas se envió una de cada suelo a cada uno de los veintiún laboratorios participantes, junto con instrucciones y un formato para remisión de resultados. En todo el proceso se veló especialmente por la confidencialidad de los resultados de cada participante. Recibidas las respuestas se procedió al análisis de resultados, asignación de valores de consenso y evaluación de cada resultado mediante el indicador “z-score”. El ejercicio ha sido muy bien recibido en un colectivo de laboratorios que no mantenía apenas contacto, y se ha mostrado muy eficaz en la caracterización de la calidad de los resultados emitidos. Se ha puesto en evidencia que hay parámetros con resultados muy homogéneos (pH) frente a otros con diferencias incompatibles con el uso agronómico fiable de los análisis (granulometría, fósforo asimilable, etc.). Queda de manifiesto la urgente necesidad de mantener periódicamente este tipo de ejercicio, como un control de calidad externo que aumente la fiabilidad de los laboratorios.

ABSTRACT

This work includes the methodology for carrying out a proficiency testing program between Spanish laboratories that offer soil fertility analysis. Aliquots of each of the two large volume samples were prepared, taken from two surface horizons of soils with marked differences between them. Once the homogeneity of aliquots was tested, one sample of each soil was sent to each of the twenty-one participating laboratories, together with the instructions and a format for the submission of results. Throughout the process, special care was taken to ensure the confidentiality of the results of each individual participant. Once the responses were received, the results were analyzed, consensus values were assigned and each result was evaluated using the “z-score” indicator. The exercise was very well received by the participating laboratories, and has been very effective in characterizing the quality of the results issued. It has been shown that there are parameters with very homogeneous results (pH) compared to others with differences incompatible with the reliable agronomic use of analyses (granulometry, assimilable phosphorus, etc.). The urgent need to maintain this type of intercomparison exercise periodically, as an external quality control that increases the reliability of laboratories, is evident.

DOI: 10.3232/SJSS.2020.V10.N3.01

RESUMO

Este trabalho inclui a metodologia para a realização de um programa de testes de proficiência entre laboratórios espanhóis que oferecem análises de fertilidade do solo. Prepararam-se alíquotas de cada uma de duas amostras de grande volume, colhidas em horizontes superficiais de dois solos, com marcadas diferenças entre si. Uma vez testada a homogeneidade das alíquotas, uma amostra de cada solo foi enviada a cada um dos vinte e um laboratórios participantes, juntamente com as instruções e um formato para a submissão dos resultados. Durante todo o processo, foram tomados cuidados especiais para garantir a confidencialidade dos resultados de cada participante. Uma vez recebidas as respostas, os resultados foram analisados, foram atribuídos valores de consenso e cada resultado foi avaliado usando o indicador "z-score". O exercício foi muito bem recebido pelo grupo de laboratórios participantes, que dificilmente mantiveram contato entre si e tem sido muito eficaz na caracterização da qualidade dos resultados emitidos. Mostrou-se que existem parâmetros com resultados muito homogêneos (pH) em comparação com outros com diferenças incompatíveis com o uso agronômico confiável das análises (granulometria, fósforo assimilável, etc.). É evidente a necessidade urgente de manter periodicamente este tipo de exercício de comparação, como um controle de qualidade externo que aumente a fiabilidade dos laboratórios.

1. Introducción y Objetivos

Los ejercicios de intercomparación constituyen una herramienta muy utilizada en laboratorios de todo el mundo para el control de la calidad (Nogueira et al. 2001). De hecho, la mayor parte de las normas de acreditación o de certificación de procedimientos en laboratorio exigen la participación en este tipo de ejercicios, lo que habitualmente suele ser costoso. Tanto la oferta comercial como los estudios de metodología para estos ejercicios son abundantes, pero suelen ir dirigidos al análisis de alimentos, presencia de drogas en diferentes matrices, metales pesados, etc., tanto de ámbito nacional como internacional (Robijns et al. 2019; Romanyukha et al. 2016; Wojciechowski et al. 2016). Pero no es fácil encontrar trabajos sobre intercomparación de analíticas de laboratorios referentes a la fertilidad de los suelos. En Europa los dos ejercicios de intercomparación con mayor oferta de parámetros de fertilidad de suelos son el programa ISE de WEPAL (Universidad de Wageningen) y el BIPEA (BIPEA 2018).

La dificultad para estos ejercicios en el ámbito de los laboratorios de fertilidad de suelos es aún mayor en procedimientos cuya aplicación requiere cantidades de muestra elevadas, o cuando se trata de parámetros para los que la oferta es muy escasa o muy específica de algunas regiones. En concreto, la determinación de la salinidad en extracto de pasta saturada, la retención de humedad o la caliza activa son ejemplos de esas situaciones.

Por otra parte, así como en otras materias (fertilizantes, piensos, grasas, residuos de pesticidas, etc.) existen en España reuniones periódicas de laboratorios en grupos de trabajo de comparación de métodos o de resultados, no existe ese tipo de intercambios en cuanto a fertilidad de suelos desde que dejó de ser operativa la Comisión de métodos oficiales del Ministerio de Agricultura, Pesca y Alimentación.

En ese contexto, y con un escaso conocimiento sobre la homogeneidad de metodologías utilizadas y de los resultados en la oferta de laboratorios españoles sobre fertilidad de suelos, se ha planteado este trabajo con el objetivo principal de establecer la metodología para realizar periódicamente ejercicios de intercomparación de análisis de fertilidad de suelos. La validación de esa metodología permitirá, además, extraer unos primeros resultados sobre el estado de la calidad de la información proporcionada por los laboratorios.

PALABRAS CLAVE
Nutrientes, laboratorios de análisis.

KEYWORDS
Nutrients, analysis laboratories.

PALAVRAS-CHAVE
Nutrientes, laboratórios de análise.

2. Material y Métodos

El primer paso para iniciar un ejercicio de intercomparación es contar muestras de suelo homogéneas. Dada la diversidad de tipos de suelos y de casuísticas de metodologías e interferencias, lo ideal sería intercambiar un número elevado de muestras diferentes, pero eso supone una dificultad muy grande al tener que preparar estas muestras estándar, y un trabajo de análisis considerable, por lo que consumiría mucho tiempo y tendría unos costes elevados. Adoptando un compromiso entre coste y beneficio, en este caso se tomaron muestras de dos suelos cultivados que presentan una respuesta del cultivo muy diferenciada. Se buscaban diferencias en el contenido de nutrientes y en parámetros como la textura o la retención de agua, y se seleccionó un suelo que tuviese problemas de salinidad.

Elegidos los dos suelos, ambos Xerofluvent típico según Keys to Soil Taxonomy (Soil Survey Staff 2014), se efectuó la toma de una cantidad de muestra suficiente (aproximadamente 80 kg), del horizonte superficial de cada uno. Las muestras fueron objeto de una primera homogeneización y disgregación en campo. Posteriormente, ya en el laboratorio, las muestras se secaron y se tamizaron para separar la fracción de tierra fina. Las muestras ya tamizadas se homogeneizaron de nuevo, esta vez en una hormigonera y, posteriormente, se prepararon cuarenta alícuotas, cada una de aproximadamente setecientos cincuenta gramos (750 g). Los envases de plástico con tapa hermética, se identificaron con una denominación de la muestra (“Huerto” y “Juntas”) y una numeración correlativa para cada una de las alícuotas.

Sobre un número representativo de esas alícuotas el laboratorio coordinador realizó un análisis de homogeneidad de todos los parámetros que se pensaban solicitar en el ejercicio, según los criterios de Olivares et al. (2018). De esas mismas alícuotas se proyectaba realizar el análisis de estabilidad temporal, mediante la nueva determinación de los parámetros susceptibles de evolucionar durante el tiempo de ejecución del ejercicio.

Y esto se hizo, en el laboratorio coordinador, una vez recibidos los resultados de todos los laboratorios.

Las analíticas a realizar debían ser lo más completas posible dentro de la oferta habitual de los laboratorios participantes (no se pedía que un laboratorio pusiese a punto un procedimiento con ocasión del ejercicio, ni que trabajase para el ejercicio de forma diferente a la habitual, ya que ambas situaciones merman la utilidad del ejercicio). Por tanto, el ejercicio estaba abierto a la participación de cada laboratorio en los parámetros que desease, de entre un máximo que incluía: pH, conductividad eléctrica en prueba previa de salinidad, materia orgánica, fósforo asimilable, potasio y magnesio asimilables, textura, microelementos y retención de agua a 33 y a 1500 kPa, para ambas muestras. Y, además, estudio de salinidad en extracto de pasta saturada, con análisis de conductividad, humedad de saturación, cationes (calcio, magnesio y sodio), y aniones (carbonatos y bicarbonatos, cloruros y sulfatos), para la que se conocía que tenía salinidad de acuerdo con el resultado de la prueba previa.

En paralelo con la preparación de las alícuotas se trabajó en la elaboración de una hoja Excel para el registro y envío de resultados de cada laboratorio. Debía ser abierta a diferentes métodos, pero cerrada en cuanto a los resultados admisibles (número de decimales y unidades, por ejemplo), y protegida frente a modificaciones que pudiese introducir un usuario concreto de forma intencionada o accidental.

Una vez encuestados hasta treinta y seis laboratorios sobre su deseo de participar, se trabajó con veintiuno, a los que se envió el siguiente material:

- Una alícuota, de unos 750 g, de cada muestra de suelos.
- Instrucciones para la realización del ejercicio.
- Hoja Excel para el envío de los resultados, con las unidades de medida para cada parámetro y con espacios para indicar la metodología utilizada.

Se codificaron las muestras y los laboratorios para asegurar la confidencialidad de la

identificación de cada laboratorio. Sólo dos responsables del laboratorio coordinador del ejercicio conocían la relación entre alícuota y laboratorio.

Sobre los resultados recibidos de los laboratorios se realizó un test de estadística robusta (Laso & Peris 2009a) para seleccionar el conjunto de valores (comprendidos en un rango de valores aceptables para obtener una desviación estándar máxima) que permitían obtener el valor que, en cada parámetro, podía considerarse referencia (o valor cierto de la propiedad), que es lo que en este tipo de ejercicios se denomina “valor de consenso”. Este valor “de consenso” es la forma habitual de otorgar un valor de referencia para una propiedad a una muestra en la que ese valor se desconoce. Conocido ese valor de consenso, que se toma en adelante como valor cierto, se calculó el indicador z-score para cada laboratorio en cada uno de los parámetros analizados (Laso & Peris 2009b). Con este análisis estadístico se clasificaron los resultados en satisfactorios ($z < 2$), cuestionables ($2 \leq z \leq 3$), o insatisfactorios ($z > 3$).

$$z\text{-score} = (V_{\text{laboratorio}} - V_{\text{consenso}}) / \text{Desviación estándar}$$

Donde:

- $V_{\text{laboratorio}}$ es el resultado emitido por el laboratorio de que se trate en el parámetro correspondiente.
- V_{consenso} es el valor de consenso, la media de los resultados de un parámetro que pasan el test de estadística robusta.
- Desviación estándar es la desviación estándar de los resultados que forman parte de la media de consenso para un parámetro.

Finalmente, el informe con el análisis estadístico de los resultados obtenidos en el ejercicio se remitió a los participantes y posteriormente se celebró una reunión de puesta en común.

3. Resultados y Discusión

Los resultados más relevantes obtenidos se muestran en la **Figura 1**. El pH(a) fue el parámetro mejor evaluado ya que hubo respuesta de todos los laboratorios y únicamente un resultado resultó cuestionable (z-score entre 2 y 3). El coeficiente de variación entre el conjunto de resultados de pH fue de en torno al 2% en ambas muestras.

En la cuantificación del contenido de materia orgánica del suelo (b) hubo más dispersión de resultados (coeficiente de variación del 57% en la muestra “Juntas”, y el 63% en la muestra “Huerto”), como se muestra en la **Figura 1**, donde se ve que tres laboratorios presentan resultados claramente insatisfactorios según el criterio de evaluación con el z-score. En los tres casos se trata de estimaciones por exceso, y en las dos muestras, lo que indica que es un error sistemático de esos laboratorios, o que el procedimiento aplicado otorga resultados no comparables. En cualquier caso, se requerirá un estudio que determine las causas. Hay que tener en cuenta que se solicitó la materia orgánica oxidable, pero en un caso se realizó la materia orgánica total por combustión, lo que permite entender la desviación de esos resultados en relación con los obtenidos por los restantes laboratorios.

El análisis granulométrico también ha dado resultados poco homogéneos (coeficiente de variación del 67% en los resultados de arena, 31% en limo, y 30% en arcilla, en la muestra “Juntas”, y del 23% en arena, 34% en limo y 25% en arcilla, en la muestra “Huerto”). La mayor parte de los laboratorios utilizan el método de sedimentación discontinua, que tiene una alta sensibilidad a las variaciones de temperatura durante el análisis (Porta et al. 1986). Aunque aparentemente el procedimiento es el mismo, es posible que haya diferencias en la fracción tamizada, en la dispersión o no de partículas, o en la destrucción o no de materia orgánica. En la reunión final se discutió sobre estos aspectos y, a partir de allí, cada laboratorio debe aplicar estos resultados y seguir participando en ejercicios de intercomparación para ajustar su metodología. En la **Figura 1(c)** se muestran los resultados

de z-score para el porcentaje de arena. De los diecinueve laboratorios participantes, once presentan resultados insatisfactorios al menos para una de las dos muestras. Hay que tener en cuenta que el rango de valores aceptables es bastante estrecho $38,27 \pm 3,54$ para la muestra "Huerto" y $14,26 \pm 1,50$ para "Juntas", y que la respuesta de los cultivos a las diferencias de contenido de arena dentro del rango aceptable en cada caso es, en la práctica, inapreciable.

Se muestran los valores de caliza activa (**Figura 1d**) por ser un parámetro muy influyente en la respuesta de los cultivos, especialmente frutales (Abadía et al. 2011), y también por ser una problemática mucho más patente en entornos semiáridos carbonatados. Se recibieron respuestas de quince laboratorios, lo que muestra que es un protocolo que no está establecido en aquellos ambientes en los que este parámetro no resulta problemático, bien por tratarse de zonas más húmedas o bien por ser terrenos más silíceos. Tres laboratorios muestran resultados insatisfactorios, dos de ellos por defecto y uno por exceso, pero en los tres casos las dos muestras presentan el mismo patrón. Hay que decir que los participantes han utilizado diferentes métodos para determinar este parámetro (se ha informado de seis métodos diferentes), y es posible que no todos sean igualmente apropiados. Cada laboratorio debe aplicar esta información para estudiar su caso.

La evaluación del contenido en fósforo asimilable por los diferentes laboratorios era uno de los principales objetivos de este ejercicio de intercomparación dirigido a parámetros de fertilidad de los suelos. Como en la determinación de la mayor parte de los nutrientes asimilables, el análisis requiere en primer lugar realizar la extracción adecuada y después cuantificar el nutriente. Los resultados de los veintidós participantes (**Figura 1e**) muestran que seis laboratorios presentan resultados insatisfactorios, tres de ellos tienen resultados concordantes para las dos muestras, pero otros tres muestran diferencias de aceptabilidad de resultados para cada una de las muestras. Se han utilizado al menos, ya que algunos laboratorios no lo han especificado, tres extractantes diferentes (Olsen, Melich III y Burriel Hernando), y al menos dos técnicas de cuantificación

(colorimetría y espectrofotometría). Esta heterogeneidad de métodos de análisis del fósforo permite explicar al menos una parte de la dispersión de resultados. El coeficiente de variación de los veintidós resultados recibidos es del 38% en la muestra "Juntas", y del 53% en la muestra "Huerto".

El extracto de pasta saturada solamente se había solicitado en la muestra "Juntas" (tenía una CE1:5 de $4,25 \pm 0,43$ dS/m) y han respondido diez laboratorios (**Figura 1f**). Excepto un laboratorio, que tiene un valor cuestionable, el resto son resultados satisfactorios. Sin embargo, debido a la gran dispersión de partida (coeficiente de variación del 47%), el rango de valores estadísticamente aceptables es muy amplio ($21,64 \pm 15,76$ dS/m a 25 °C), tanto que engloba desde valores que no afectarían a muchos cultivos hasta otros que hacen imposible cualquier tipo de vegetación, lo que indica que, dado que el número de respuestas obtenidas ha sido muy bajo, el test estadístico utilizado no cuenta con suficientes datos.

4. Conclusiones

Del estudio realizado se pueden extraer las conclusiones siguientes:

- El ejercicio diseñado ha sido eficaz para concitar interés de diversos laboratorios y obtener, en gran parte de los parámetros, un número de respuestas que permite el análisis. La información obtenida ha sido percibida muy positivamente por los laboratorios participantes por su utilidad en la homogeneización de metodología y resultados.
- La metodología puesta a punto para ejercicios de intercomparación ha resultado de interés, eficaz, robusta, y es reproducible periódicamente.
- El análisis realizado pone de manifiesto la falta de coherencia de resultados analíticos y, por ende, la necesidad de utilizar herramientas para reducir su heterogeneidad. Y el ejercicio de intercomparación se manifiesta como la más eficaz.

- Las diferencias observadas en los métodos de análisis utilizados por distintos laboratorios en España evidencia la necesidad de crear un ente coordinador, como podría ser un

laboratorio de referencia de suelos, tal como propone el GLOSOLAN, impulsado por el Ministerio de Agricultura, Pesca y Alimentación.

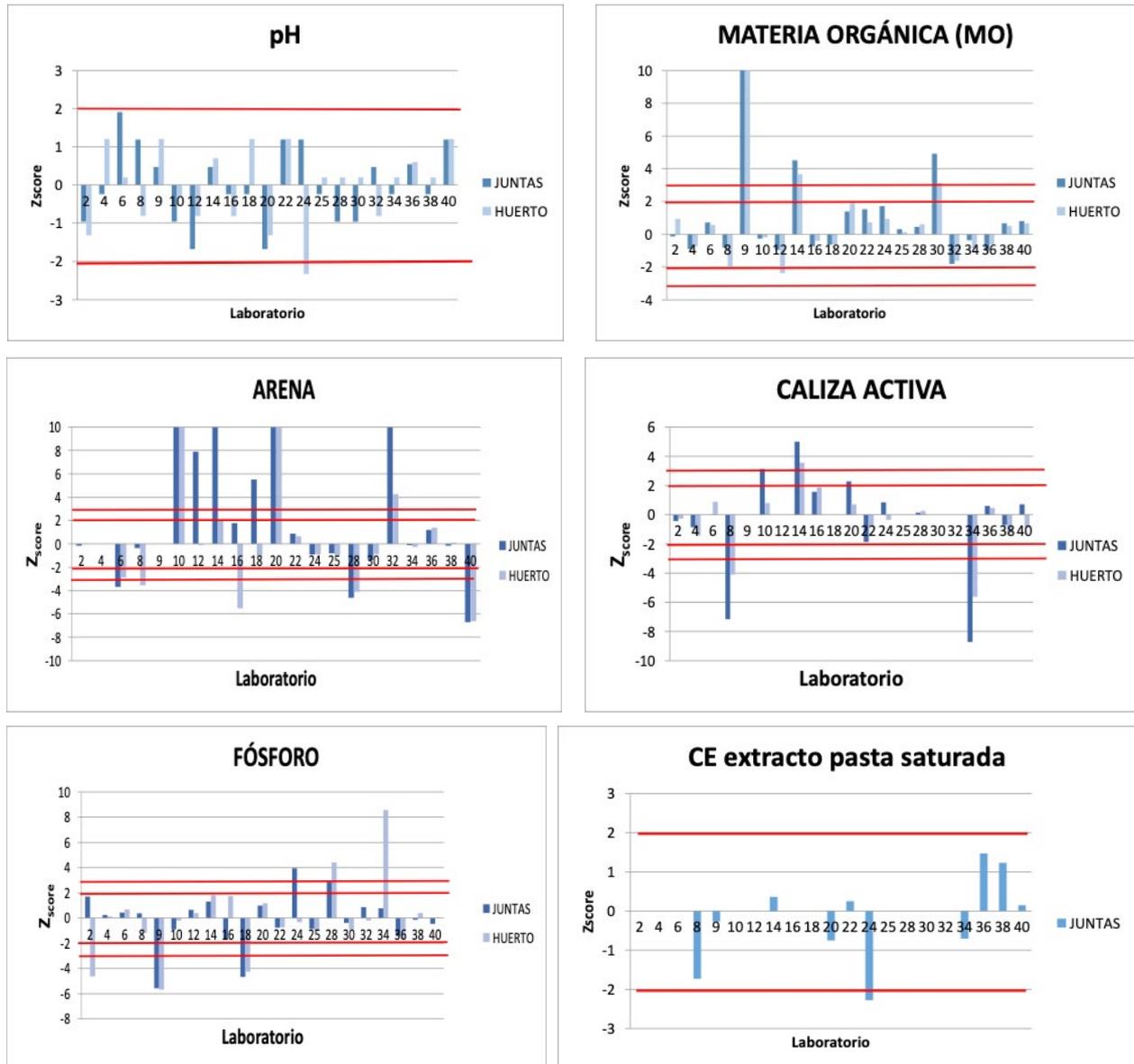


Figura 1. Resultados de z-score para algunos de los parámetros analizados. (a) pH, respuestas de 21 laboratorios, n consenso de 21; (b) Materia orgánica, datos expresados en %, respuestas de 21 laboratorios, n consenso de 15 en “Huerto” y 17 en “Juntas”; (c) Arena, datos expresados en %, respuestas de 19 laboratorios, n consenso de 12 en “Huerto” y 10 en “Juntas”; (d) Caliza activa, datos expresados en %, respuestas de 15 laboratorios, n consenso de 12 en “Huerto” y 10 en “Juntas”; (e) Fósforo asimilable, datos expresados en mg/kg, respuestas de 21 laboratorios, n consenso de 16; (f) CE del extracto de pasta satura para la muestra “Juntas”, datos expresados en dS/m, respuestas de diez laboratorios, n consenso de nueve.

5. Agradecimientos

Los autores desean manifestar su agradecimiento al Ministerio de Agricultura, Pesca y Alimentación, que ha apoyado el proyecto; a la Sociedad Española de la Ciencia del Suelo, cuyo respaldo ha aportado solvencia a este proyecto y probablemente también le facilite la continuidad; y a la Unidad de Suelos del Laboratorio Agroambiental, del Gobierno de Aragón, que ha prestado apoyo analítico y logístico fundamental.

REFERENCIAS

- Abadía JS, Vázquez R, Rellán-Álvarez H, El-Jendoubi A, Abadía A, Álvarez-Fernández A, López-Millán F. 2011. Towards a knowledge-based correction of iron chlorosis. *Plant Physiology and Biochemistry* 49(5):471-482.
- BIPEA. 2018. International Proficiency test provider. Disponible en: <https://extranet.bipea.org/documents/plannings/GB/Proficiency-testing-scheme-15.pdf> extraído el 9/10/2019.
- Laso J, Peris A. 2009a. Evaluación de resultados de ensayos de aptitud. En: *Comunicaciones del V Iberolab*; Madrid; p. 1-5.
- Laso J, Peris A. 2009b. Tratamientos estadísticos de aptitud: aplicación de la mediana para detección de resultados anómalos. En: *Comunicaciones del V Iberolab*; Madrid; p. 6-11,
- Nogueira JM, Nieto de Castro F. 2001. EPTIS: The new European database of proficiency testing schemes for analytical labs. *Trends in Analytical Chemistry* 20(9):457-461.
- Olivares IRB, Souza GB, Nogueira ARA, Toledo GTK, Marcki DC. 2018. Trends in developments of certified reference materials. *Trends in Analytical Chemistry* 100:53-64.
- Porta J, López-Acevedo M, Rodríguez R. 1986. *Técnicas y experimentos en edafología*. Colegi Oficial d'Enginyers Agrònoms de Catalunya. 282 p.
- Robijns K, Boone NW, Jansen RTP, Kuypers AWHM, Neef C, Touw DJ. 2019. Commutability of proficiency testing material containing amitriptyline and nortriptyline: A study within the framework of the Dutch Calibration 2.000 project. *Clinica Chimica Acta* 498:6-10.
- Romanyukha A, Grypp MD, Fairchild GR, Williams AS. 2016. Performance comparison of OSLD (Al₂O₃:C) and TLD (LiF: Mg, Cu, P) in accreditation proficiency testing. *Radiation Measurements* 93:7-12.
- Soil Survey Staff. 2014. *Keys to Soil Taxonomy*. 12th ed. Washington, D. C.: USDA-Natural Resources Conservation Service.
- WEPAL. 2018. International Soil- Analytical Exchange. Quarterly Report. University of Wageningen. Disponible en: <http://www.wepal.nl/> extraído el 9/10/2019.
- Wojciechowski KL, Melilli C, Barbano DM. 2016. A proficiency test system to improve performance of milk analysis methods and produce reference values for component calibration samples for infrared milk analysis. *Journal of Dairy Science* 99(8):6808-6827.