UNIVERSIDAD TÉCNICA DE MANABÍ
Fundada en 1952

LA TÉCNICA
Revista de las Agrociencias

Agricultura y Silvicultura

Importance of joint interpretation of the coefficient of variation and determination in agricultural trials: A case study with bean yield

Importancia de la interpretación conjunta del coeficiente de variación y de determinación en ensayos agrícolas: Un estudio de caso con rendimiento de frijol

Authors

Freddy Carlos Gavilánez Luna
fgavilanez@uagraria.edu.ec

Universidad Agraria del Ecuador, Guayaquil, Ecuador.

Suggested citation: Gavilánez Luna, F. C. (2025). IImportance of joint interpretation of the coefficient of variation and determination in agricultural trials: A case study with bean yield. *La Técnica, 15*(2), 97-102. DOI: https://doi.org/10.33936/latecnica.v15i2.7320

Received: February 25th, 2025
Accepted: September 22nd, 2025
Published: October 25th, 2025

Abstract

Verifying the goodness of an agricultural experiment in analysis of variance (ANOVA) models only through the coefficient of variation (CV) of its variables can lead us to make decisions about imprecise trials (error II) or even take us to the extreme. more complicated to recommend alternatives that in practice do not have significant effects (error I). A statistician that must accompany the CV to assess this goodness and that is not affected by the degrees of freedom of the experimental error as in the latter, is the coefficient of determination (r^2). Given this situation, it was proposed to demonstratively detail a practical example to increase the value judgment of a researcher when accepting the results of an experiment, considering for this purpose what is established by the CV and the r^2 simultaneously. Performance was analyzed as a variable in an experiment with 24 data, considering a randomized complete block design with six treatments and four repetitions. The assumptions of normality, independence and homoscedasticity of the residuals were checked; the ANOVA was performed and the CV, r^2, was calculated, in addition to the repeatability index of the trial. The apparent goodness of the experiment was demonstrated by obtaining a CV of 10.88%, well below the maximum limit given by the literature (30%), although the r^2 barely presented a value of 0.47 far from the minimum limit of 0.60; which highlights the importance of looking at CV alongside r^2 when examining the robustness of an experimental result.

Keywords: analysis of variance, experimental error, experimental unit, sums of squares.

Resumen

Verificar la bondad de un experimento agrícola en los modelos de análisis de varianza (ANOVA) únicamente a través del coeficiente de variación (CV) de sus variables puede inducir a tomar decisiones sobre ensayos poco precisos (error II) o, incluso, llevar al extremo más complicado de recomendar alternativas que en la práctica no tienen efectos significativos (error I). Un estadígrafo que debe acompañar al CV para valorar dicha bondad y que no sea afectado por los grados de libertad del error experimental como en este último, es el coeficiente de determinación (r^2). Ante esta situación se propuso detallar demostrativamente un ejemplo práctico para incrementar el juicio de valor de un investigador al momento de aceptar los resultados de un experimento, considerando para ello lo establecido por el CV y el r^2 de manera simultánea. Se analizó el rendimiento como variable de un experimento con 24 datos, considerando un diseño de bloques completos al azar con seis tratamientos y cuatro repeticiones. Se comprobaron los supuestos de normalidad, independencia y homocedasticidad de los residuos; se realizó el ANOVA y se calculó el CV, r^2 y el índice de repetitividad del ensayo. Se demostró la aparente bondad del experimento al obtenerse un CV de 10,88%, por debajo del límite máximo dado por la literatura (30%); no obstante, que el r^2 apenas presentó un valor de 0,47 alejado del límite mínimo de 0,60; lo cual destaca la importancia de observar el CV junto al r^2 cuando se examina la solidez de un resultado experimental.

Palabras clave: análisis de varianza, error experimental, sumas de cuadrados, unidad experimental.

Introduction

Assays are important decision-making tools in every experimental research, specially when they call for a methodological change that requires a higher investment than what was initially planned for the production process. In this context, experiments in agriculture are not the exception, and these, most of the time, and specially when they require external validation, serve to recommend results to agriculture workers from an area, so they adopt the suggested changes, for example, implementing new seeds, a different dosage or fertilizer, a new way of treating the soil, a complementary application of organic nutrients, among others; the main purpose being achieving higher profits.

Nonetheless, decision-making with regards to the results from an experiment may have chance of commitign error I, that is, ignoring a reality in which all tests have a similar statistical effect (rejecting the null hypothesis when it is true), therefore resulting in a false positive; or error II, which refers to declaring no significant differences between groups of treatments when, in reality, they have different effects (accepting the null hypothesis when it its false (Kuehl, 2001), resulting in a false negative. The only way to avoid the former is randomizing the treatment assignments to the experimental unit or vice versa; meanwhile, the latter may be avoided by increasing the number of repetitions (Gutiérrez y De La Vara, 2008).

Error II is linked to experimental accuracy and it occurs when the experimental error reachers a certain value, and in that case, a common statistical tool such as analysis of variance (ANOVA) does not dectect significant differences between treatments, even if they do exist (Vargas-Rojas et al., 2020). Nontheless, the variance or mean square of the experimental error (MSE) can be minimized by increasing the number of replications, thereby achieving an apparent sensitivity in the experiment to detect differences, illusorily reducing even the effect of any possible systematic error that may have been added to the experimental error, even though the latter should be purely random (Cochran and Cox, 1992).Most agricultural researchers use the MSE to assess the soundness of their experiments through the coefficient of variation (CV) (Nardino et al., 2020), even though there are critics of the use of the latter due to its dependence on the means (Taylor et al., 2008; Vásquez and Caballero, 2011). However, the CV offers an approximate view of the degree of care with which a trial has been conducted (Balzarini et al., 2011), and values above 30% are considered an indication of unreliable experiments (Mićić and Bosančić, 2012; Gordón-Mendoza and Camargo-Buitrago, 2015) when their ANOVA models yield significant differences among treatments. In this sense, increasing the number of replications can cause the CV to decrease, even to the point of declaring "significant" differences among treatments that are impracticable in terms of profitability and that do not even exceed the producer's opportunity costs (CIMMYT, 1988), especially when a CV lower than 5% is recorded (Mićić and Bosančić, 2012).

In a poorly managed experiment, a relatively high number of replications will only mathematically minimize the systematic error produced (which is not part of the experimental error); nevertheless, this error will remain present in the trial. Therefore, making a decision under this circumstance can be risky. This risk is insufficiently warned of by the CV because this coefficient is calculated from the MSE, that is, a value obtained from the division between the quadratic residuals of the experimental error (SSE) of the data and their corresponding degrees of freedom (Montgomery, 2004), which depend on the number of replications, ultimately masking the real variability within each treatment (intra-group variability) by presenting a supposedly acceptable CV (less than 30%) (Gordón-Mendoza and Camargo-Buitrago, 2015).

A statistic that must inevitably be observed in a complementary manner to the CV in order to assess the precision of an experiment and to detect the aforementioned shortcomings is the coefficient of determination (r^2). This coefficient, unlike the CV, takes into account the absolute variation of the data represented by the sums of squares, indicating the proportion of variability that is due to the factors of interest evaluated within a given ANOVA model (explained variability) with respect to the total variability (Montgomery et al., 2006). In this sense, the r^2 of an ANOVA model, which can range between 0 and 1, makes it possible to

latecnica@utm.edu.ec | Vol. 15, Num. 2 (97-102): July-December, 2025 | DOI: 10.33936/latecnica.v15i2.7320

observe the absolute variation (unaffected by degrees of freedom) that corresponds to the experimental error ($1 - r^2$) and to have greater certainty about the soundness of an experiment when its value is ≥ 0.7 (Gutiérrez and De La Vara, 2008), or even if a lower limit such as 0.6 is considered (Balzarini et al., 2011).

The coefficient of determination is one of the most widely used measures of fit for linear models, whether they pursue prospective or confirmatory objectives (García, 2002). Comparatively, r^2 is more immutable than the CV in the sense that its values are not relativized according to sample size (Chicco et al., 2021), and its greater informative capacity can be observed in the selection of the levels (treatments) of the test factors in experiments (Gutiérrez and De La Vara, 2008), especially those of a quantitative nature such as fertilizers, since poorly selected doses could lead to low explained variability, resulting in unacceptable r^2 values (<0.6) (Balzarini et al., 2011). In addition, one noteworthy aspect of r^2 is its ability to remain unchanged even when certain data transformations are applied to stabilize variances (Bowman and Watson, 1997), thus constituting a fundamental element when declaring significant effects

Therefore, the objective of this technical note is to strengthen a researcher's value judgment when accepting the results of an experiment, by simultaneously considering what is established by the CV and r^2..

Materials and methods

Yield data of canary bean (Phaseolus vulgaris L.), adjusted to 12% moisture, were used from one of the practice experiments conducted at the Experimental Center "El Misionero" of the Universidad Agraria del Ecuador (Table 1) in the city of Milagro, Guayas Province. The climatic conditions of the area are characterized by a mean annual precipitation of 1,297 mm, a relative humidity of 82%, an average temperature of 25.2 °C, and a wind speed of 1.2 $m \cdot s^{-1}$. The trial was carried out as part of the practical component of the experimental design course of the agronomy program, in which increasing doses of the leachate from a vermiculture process were evaluated (from 0 to 25 $L \cdot ha^{-1}$ at intervals of 5 $L \cdot ha^{-1}$). The trial included six treatments, each with four replications, under a randomized complete block design. In order to ensure the statistical reliability of the experiment, it was designed so that the experimental error had a minimum of 12 degrees of freedom, in accordance with the criterion proposed by Carballo and Quiroga (1976).

Table 1. Yield data ($kg \cdot ha^{-1}$) of canary beans in Milagro, Ecuador 2024.

Blocks	Treatment						Total
	T1 (0 $L \cdot ha^{-1}$)	T2 (5 $L \cdot ha^{-1}$)	T3 (10 $L \cdot ha^{-1}$)	T4 (15 $L \cdot ha^{-1}$)	T5 (20 $L \cdot ha^{-1}$)	T6 (25 $L \cdot ha^{-1}$)	
1	990.90	1241.4	1315.9	1368.9	1269.9	1080.6	7267.6
2	860.10	1181.0	1287.3	1148.6	1218.7	1109.5	6805.2
3	1350.2	971.0	1132.1	1311.2	1154.9	1027.3	6946.7
4	957.10	1103.1	1336.7	1222.9	1351.2	1184.5	7155.5
Total	4158,3	4496.5	5072.0	5051.6	4994.7	4401.9	28175.0
Standard Deviation	214,38	116.73	92.84	97.01	82.97	65.55	
Medias	1039,6	1124.1	1268.0	1262.9	1248.7	1100.5	

With the information, the ANOVA was performed, after verifying the fulfillment of the assumptions of normality and homoscedasticity of the residuals through analytical and graphical methods. At the same time, the CV, r^2, and the repeatability index (Ir) were obtained. These analyses were carried out using the RStudio software (RStudio Team, 2023).

Results and discussion

The residuals of the data practically met the assumptions of normality, homoscedasticity, and independence (Figure 1), all of which made it possible to validate the results subsequently obtained through the ANOVA model (Montgomery, 2004), even before the incidence of the extreme value represented in treatment 1 of block 3 became evident through the boxplot, as well as through the homoscedasticity and independence plots of the residuals. The incidence of this single value did not alter the level of significance ($P > 0.05$) revealed by the Shapiro–Wilk and Levene tests, which respectively verified the fulfillment of normality and homoscedasticity of the residuals (table 2).

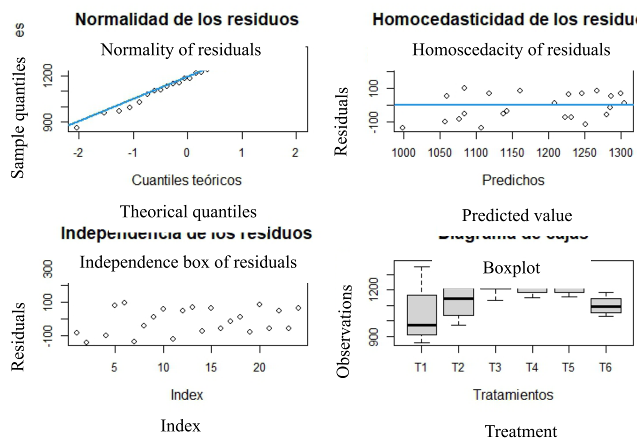

Figure 1. Diagnostic plots for checking ANOVA assumptions.

Analytical tests for exploratory analysis of experimental data, given the sample sizes typically handled in such studies, must be applied in conjunction with graphical evaluations in order to effectively verify the fulfillment of the assumptions of normality and homoscedasticity of the residuals. In this case, if only the analytical perspective is considered, the fulfillment of normality and homoscedasticity of the data would have to be taken for granted (table 2); however, the graphical analysis made an extreme value among the data evident. Hence, a certain preference is given to the latter in experiments, since it makes it possible to observe whether such noncompliance is supported by more than two outliers, which is the limit for irreversibly restricting the use of parametric tests (Gutiérrez and De La Vara, 2008).

Table 2. Analytical verification of the assumptions of normality and homoscedasticity of yield data (kg·ha⁻¹) of canary bean in Milagro, Ecuador, 2024.

	Observed statistic	*P*-value
Shapiro - Wilk	0.9543	0.3355
Levene[1]	1.4303	0.2610

[1]Centered on the mean.

The ANOVA of the data revealed non-significant differences (0.090) among treatments (table 3) at the 5% probability level. However, the differences could be declared significant if a 10% error probability were taken as the threshold; a different situation applies to the blocks, which revealed a clear homogeneity of the field (P-value= 0.729). This apparent certainty is supported by the CV of the data, as it is below the maximum 30% recommended for agricultural experiments (Balzarini et al., 2011). However, when the sums of squares of the three sources of variation of the ANOVA are examined, the greater variation presented by the experimental error (53.3%) with respect to that of treatments and blocks is noteworthy, evidencing a substantial load of "unknown" variance that renders the trial insensitive and prevents it from detecting significant differences.

Table 3. ANOVA of yield data (kg·ha⁻¹) of canary bean in Milagro, Ecuador, 2024.

Sources of variation	GL	SC	CM	F	*P*-value
Total	23	459498.52			
Treatments	5	193116.31	38623.26	2.366	0.090
Blocks	3	21486.95	7162.32	0.439	0.729
Experimental error	15	244895.26	16326.35		

CV: 10.88%, r^2: 0.47, I_r: 0.47.

The magnitude of the experimental error indicated in table 3 is fully highlighted when observing the coefficient of determination of the trial, with a value lower than the minimum recommended 0.70 required to have confidence in what is resolved by an experiment in terms of precision (Gutiérrez and De La Vara, 2008). Complementarily, this precision can also be observed through the repeatability index (Ir) of the experiment, whose value, being below 0.70, is likewise evidence of the lack of care in the trial (Senar, 1999).

When observing the CV value, the trial would apparently be considered to have good precision and uniformity, since this value was lower than 20% (Ruiz-Ramírez, 2010; Mishra et al., 2023). However, the variability shown by this indicator is somewhat misleading when the r^2 of the trial is examined simultaneously, which reveals that the variation actually explained by the ANOVA model is barely 47%; that is, an experiment that lacks reliability (Balzarini et al., 2011) due to the substantial unknown absolute variability that detracts from any result. At this point, it should be noted that, among the various agricultural trials, experiments with relatively high CV values do not necessarily imply the notable presence of systematic error, since this coefficient depends on the type of experiment and on the class of variable being evaluated (García et al., 2021); however, yield is a variable that commonly presents normal residuals due to the internal control of the trials (Yaguas, 2017).

One defective aspect masked by the CV in experiments is related to the degrees of freedom of the experimental error (DFEE). Its apparent soundness gains strength from approximately 12 to 14 DFEE, which constitute approximate lower limits from which several authors (Mayor-Durán et al., 2012; Mishra et al., 2023) justify the statistical power of tests such as ANOVA to achieve experimental sensitivity. Even so, having an experiment with a high number of DFEE makes it possible to conceal the absolute magnitude of the experimental error through a reduction of the CV, as has occurred in this experiment; this can even be counterproductive from the standpoint of committing a Type I error when the ANOVA manages to detect minimal statistical differences among treatments that, in practical application, are not convenient from the standpoint of profitability (Cumming, 2014; Martínez-Ezquerro et al., 2017), since the magnitude of the increase in crop yield may not offset the additional costs, which would instead reduce the economic profitability of the process. Moreover, the excessive increase in replications beyond what is necessary, in addition to generating expenses, does not contribute to a substantial increase in precision either (Martínez-Ezquerro et al., 2017); and, under the presence of non-random error due to poor experimental management, the latter would only be apparent precision.

latecnica@utm.edu.ec | Vol. 15, Num. 2 (97-102): July-December, 2025 | DOI: 10.33936/latecnica.v15i2.7320

Conclusion

The validity of an experiment in terms of precision should not be based exclusively on what the CV reveals, since this coefficient may mask the true magnitude of the experimental error. For this reason, it is necessary to consider other indicators, such as r², which make it possible to assess, in absolute and comprehensive terms, the degree of reliability and robustness of a trial.

Conflict of interest

The author declares that there are no conflicts of interest in the present publication at any of its stages.

References

Balzarini, M., Di Rienzo, J., Tablada, M., González, L., Bruno, C., Córdoba, M., Robledo, W. y Casanoves., F. (2011). *Estadística y biometría*. Argentina: Editorial Brujas.

Bowman, D. T. and Watson, C. E. (1997). Measures of validity in cultivar performance trials. *Agronomy Journal, 89*(6), 860-866. https://doi.org/10.2134/agronj1997.00021962008900060003x

Carballo, A. y Quiroga, V. (1976). *Manual práctico para el análisis de experimentos de campo*. Instituto Interamericano de Ciencias Agrícolas (IICA). https://n9.cl/kvk9a

Centro Internacional de mejoramiento de maíz y trigo (CIMMYT). (1988). *La formulación de recomendaciones a partir de datos agronómicos: Un manual metodológico de evaluación económica*. México: CIMMYT. https://n9.cl/47h35

Chicco, D., Warrens, M. J. and Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*:e623. https://doi.org/10.7717/peerj-cs.623

Cochran, W. G. and Cox, G. M. (1992). *Experimental designs*. Reino Unido: Wiley.

Cumming G. (2014). The new statistics: why and how. *Psychol. Sci*., *25*(1), 7-29. doi: 10.1177/0956797613504966.

García, B., Amorin, G., Luciene, K., Suares, P., Bastos, A., Almeida, T., Daieny, A. and Menezes, L. (2021). Classification of the coefficient of variation for experiments with eucalyptus seedlings in greenhouse. *Revista Ciência Agronômica, 52*(4), e20207587. https://doi.org/10.5935/1806-6690.20210050

García, R. M. (2002). *Inferencia estadística y diseño de experimentos*. Buenos Aires, Argentina: EUDEBA.

Gordón-Mendoza, R. y Camargo-Buitrago, I. (2015). Selección de estadísticos para la estimación de la precisión experimental en ensayos de maíz. *Agronomía Mesoamericana, 26*(1), 55-63. http://dx.doi.org/10.15517/am.v26i1.16920

Gutiérrez, H. y De La Vara, R. (2008). *Análisis y diseño de experimentos* (2da. Ed.). México: McGraw-Hill/ Interamericana Editores S.A.

Kuehl, R. O. (2001). *Diseño de experimentos: principios estadísticos de diseño y análisis de investigación* (2da. Ed.). México: International Thomson Editores S.A.

Mayor-Durán, V., Blair, M. y Muñoz, J. (2012). Metodología para estimar el coeficiente de heterogeneidad del suelo, el número de repeticiones y el tamaño de parcela en investigaciones con frijol (*Phaseolus vulgaris* L.). *Acta Agronómica, 61*(1), 32-39. http://www.scielo.org.co/scielo.php?pid=S0120-28122012000100005&script=sci_arttext

Mićić, N. and Bosančić, B. (2012). Variability and variation coefficients in biological and agricultural experimental research. *Agroznanje, 13*(3), 331-342. https://doi.org/10.7251/AGRSR1203331M

Mishra, P., Hasham, M. M., Manjunatha, G. R., Ray, S., Homa, F. and Pawariya, V. (2023). Role of design of experiment in agriculture trials and its role in decision making. *International Journal of Agricultural & Statistical Sciences, 19*(1), 389-399. http://dx.doi.org/10.59467/IJASS.2023.19.389

Montgomery, D. C. (2004). *Diseño y análisis de experimentos* (2da. Ed.). México: Editorial LIMUSA S.A.

Montgomery, D. C., Peck, E. A. y Vining, G. G. (2006). *Introducción al análisis de regresión lineal* (3era. Ed.). México: Compañía Editorial Continental.

Nardino, M., Machado, J., Torres, V., Costa, F. Dias, F. and Silva, W. (2020). Coefficient of variation: a new approach for the study in maize experiments. *Brazilian Journal of Biometrics, 38*(2), 185-206. https://doi.org/10.28951/rbb.v38i2.440

RStudio Team. (2023). *RStudio: Integrated Development for R* [Computer software]. Posit Software, PBC. https://posit.co

Ruiz-Ramírez, J. (2010). Eficiencia relativa y calidad de los experimentos de fertilización en el cultivo de caña de azúcar. *Terra Latinoamericana, 28*(2), 149-154. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-57792010000200006

Senar, J. C. (1999). La medición de la repetibilidad y el error de medida. *Etologuía, 5*(17), 53-64. https://www.researchgate.net/publication/233801237_La_medicion_de_la_repetibilidad_y_el_error_de_medida

Taylor, S. L., Payton, M. E. and Raun, W. R. (2008). Relationship between mean yield, coefficient of variation, mean square error, and plot size in wheat field experiments. *Communications in Soil Science* *and Plant Analysis, 30*(9-10), 1439-1447. https://doi.org/10.1080/00103629909370298

Vargas-Rojas, J., Vega-Villalobos, E. y Cerdas-Ramírez, R. (2020). Tamaño y forma de la unidad experimental en ensayos de rendimiento de *Brachiaria* híbrido CIAT 3608. *Pastos y Forrajes, 43*(2), 144-149. http://scielo.sld.cu/pdf/pyf/v43n2/2078-8452-pyf-43-02-144.pdf

Vásquez, E. R. y Caballero, A. (2011). Inconsistencia del coeficiente de variación para expresar la variabilidad de un experimento en un modelo de análisis de varianza. *Cultivos Tropicales, 32*(3), 59-62. https://www.redalyc.org/articulo.oa?id=193222357008

Declaración de contribución a la autoría según CRediT

Freddy Carlos Gavilánez Luna: conceptualización, redacción-borrador original, metodología, investigación, análisis formal, redacción-revisión y edición.