



## Estudio de la producción de frijol (*Phaseolus vulgaris* L.) en la región del Bajío: un análisis estadístico multivariado

**María Janet Calvillo García \***

Tecnológico Nacional de México / IT de  
Celaya,  
Celaya, Guanajuato, México

**José Alfredo Jiménez García**

Tecnológico Nacional de México / IT de  
Celaya,  
Celaya, Guanajuato, México

**David Israel Contreras Medina**

Tecnológico Nacional de México / IT de  
Celaya,  
Celaya, Guanajuato, México

\* Autor de correspondencia: [m2403098@itcelaya.edu.mx](mailto:m2403098@itcelaya.edu.mx)

---

**Resumen:** Este estudio analizó la producción de frijol de los años 2013 al 2023 mediante estadística multivariada. Se identificaron correlaciones significativas: superficie-producción  $-0.541$  a  $0.986$  y superficie-rendimiento negativas en Irapuato y Celaya. Los modelos de regresión mostraron  $R^2$  alto  $0.98$  en Irapuato y bajo  $0.10$  en San Luis de La Paz. El análisis de componentes principales reveló que 70%-95% vinculó capacidad agrícola, y un 30% vinculó áreas siniestradas. Destacando desigualdades extremas San Felipe: 10,134 ton vs. Tierra Blanca: 42 ton.

**Palabras clave:** Adopción, estadística multivariada, frijol (*Phaseolus vulgaris* L.), tecnología sostenible.

---

### 1. Introducción: lo que debemos saber de inicio

En la actualidad, la adopción de la tecnología, impulsada por tecnologías como IoT, Big Data, drones y automatización, es clave para lograr una agricultura sostenible en pequeños productores, incluidos los de frijol (*Phaseolus vulgaris* L.). Organizaciones como la

Organización de las Naciones Unidas para la Agricultura y la Alimentación (*Food and Agriculture Organization of the United Nations*, FAO) y el Foro Económico Mundial destacan su potencial en regiones con escasos recursos, ejemplificando proyectos como *Climate-Smart Agriculture*, que integra IoT para mejorar la resiliencia climática. Sin embargo, factores socioeconómicos limitan su adopción en sectores vulnerables. A nivel global, estudios analizan variedades de frijol, técnicas de cocción (Wiesinger et al., 2021) y compostaje orgánico para maximizar rendimiento (Rodríguez Espinoza et al., 2024). En México, la producción enfrenta una crisis: con una caída anual del 3.2% y dependencia de importaciones, se amenaza la seguridad alimentaria. Investigaciones como la de Servín-Palestina et al. (2022) describen este declive, pero emplean métodos univariados, sin profundizar en análisis multivariados. Contreras-medina et al. (2025) abordan la desconexión entre innovación tecnológica y conocimiento tradicional, con enfoque estadístico es limitado. Kadege et al. (2024) aplican ANOVA para estudiar resistencia a enfermedades, sin llegar a realizar un análisis con estadística multivariada. La integración exitosa de la tecnología requiere equilibrar participación comunitaria y marcos analíticos robustos que incorporen variables económicas, sociales y ambientales. Solo así se garantizará una agricultura sostenible ante el crecimiento poblacional mundial, especialmente en cultivos estratégicos como el frijol.

## **2. Fundamentos Teóricos: reglas y principios científicos importantes**

El frijol, esencial para la seguridad alimentaria en México y Latinoamérica, requiere estudios que analicen su producción mediante estadística multivariada. Esta metodología permite evaluar interacciones complejas entre variables agrícolas (clima, suelo, tecnología, prácticas) que influyen en su productividad y sostenibilidad. Técnicas como correlaciones y regresión lineal múltiple (Pardo, 2020), optimizan el uso de recursos (agua, fertilizantes) e identifican oportunidades para implementar soluciones tecnológicas, mitigando riesgos de escasez y dependencia de importaciones. En Guanajuato, región clave del Bajío, el frijol es prioritario: 836,105 hectáreas están sembradas, con municipios como Ocampo (80% de superficie agrícola) y San Felipe (50%) dedicados mayoritariamente a este cultivo (Gallegos Acosta et al., 2018).

La revisión bibliográfica indica vacíos en estudios recientes sobre producción de frijol en México que apliquen estadística multivariada. Por ejemplo, Ping et al. (2024) analizan habas con estas técnicas, mientras Zamukulu et al. (2023) estudian al frijol (*Phaseolus vulgaris* L.) usando ANOVA y pruebas de diferencia mínima significativa, sin explorar relaciones multivariadas. Esto resalta la necesidad de enfoques integrales que superen análisis univariados o bivariados. Esta investigación aplicará estadística multivariada a datos de producción de frijol en el Bajío, evaluando cómo variables como tecnología, siembra, cosecha, rendimiento y condiciones ambientales interactúan. El objetivo es determinar si la adopción tecnológica influye significativamente en la productividad, proponiendo estrategias basadas en evidencia para fortalecer la sostenibilidad agrícola y reducir vulnerabilidades en un contexto de crecimiento poblacional y cambio climático.

### **3. Desarrollo del Trabajo: aplicando las reglas y principios científicos**

Se estudiaron 43 de 46 municipios guanajuatenses con datos completos de producción de frijol (2013-2023), analizando superficie sembrada (ha), producción (ton) y rendimiento (ton/ha).

Las estadísticas descriptivas revelaron una disparidad extrema: Los 5 principales productores son San Felipe (10,134.18 ton), Ocampo (5,339.05 ton), San Luis de La Paz (3,724.70 ton), Dolores Hidalgo Cuna de la Independencia Nacional (3,360.04 ton) y Salvatierra (2,720.06 ton), mientras los menores fueron Tierra Blanca (42.61 ton), Atarjea (68.67 ton), Santa Catarina (81.28 ton), Purísima del Rincón (97.66 ton) y Cuerámara (121.49).

*Coeficiente de correlación de Pearson:*

Se utilizó el coeficiente de correlación de Pearson para medir la correlación entre dos conjuntos de datos.

- **Superficie vs Producción:** Con un resultado positivo en Irapuato (0.986), Huanímaro (0.970), Pénjamo (0.975), y Villagrán (0.976), indicando que más superficie aumenta la producción. Con excepciones como Guanajuato (-0.452), Victoria (-0.541), Atarjea (-0.082) y San José Iturbide (-0.288) indicando que aumentar la superficie no garantiza mayor producción.

- **Superficie vs Rendimiento:** Predominan correlaciones negativas en Irapuato (-0.745), Celaya (-0.614), Comonfort (-0.734), y Jerécuaro (-0.697), sugiriendo ineficiencia al expandir cultivos. Salamanca (0.313), Moroleón (0.200), y Villagrán (0.648) son excepciones donde más superficie se vincula a mayor rendimiento, lo que podría indicar prácticas agrícolas eficientes o cultivos altamente productivos.
- **Producción vs Rendimiento:** Resultados mixtos con correlaciones positivas en San Miguel de Allende (0.979), Ocampo (0.969), Dolores Hidalgo (0.364), y Xichú (0.793) mostrando que un aumento en la producción está ligado a un mayor rendimiento, mientras Irapuato (-0.644), Silao (-0.394), y Comonfort (-0.196), reflejan sobreproducción en tierras poco aptas o prácticas que degradan la productividad por hectárea.

Tabla 1. Resultados de Abasolo.

Abasolo			
Análisis de los valores de la matriz de correlación			
Valor propio	2.8668	0.1332	0.0000
Proporción	0.956	0.044	0.000
Acumulada	0.956	1.000	1.000
Variable	PC1	PC2	PC3
Siembra	0.584	0.398	-0.707
Cosecha	0.584	0.398	0.707
Producción	0.563	-0.826	0.000

Obtenida de: elaboración propia.

Tabla 2. Resultados de Cortázar

Cortázar			
Análisis de los valores de la matriz de correlación			
Valor propio	2.7715	0.2285	0.0000
Proporción	0.924	0.076	0.000
Acumulada	0.924	1.000	1.000
Variable	PC1	PC2	PC3
Siembra	0.590	0.389	-0.707
Cosecha	0.590	0.389	0.707
Producción	0.551	-0.835	0.000

Obtenida de: elaboración propia.

Tabla 3. Resultados de San Francisco del Rincón

San Francisco del Rincón			
Análisis de los valores de la matriz de correlación			
Valor propio	2.9412	0.0588	0.0000
Proporción	0.980	0.020	0.000
Acumulada	0.980	1.000	1.000
Variable	PC1	PC2	PC3
Siembra	0.580	0.404	-0.707
Cosecha	0.580	0.404	0.707
Producción	0.571	-0.821	0.000

Obtenida de: elaboración propia.

### Modelo de regresión lineal múltiple:

Se evaluó un modelo de regresión lineal múltiple. Los municipios con  $R^2$  alto como Irapuato (0.98), Pénjamo (0.96), Valle de Santiago (0.96), Villagrán (0.95), y Huanímaro (0.94) confirmaron que la superficie sembrada explica casi toda la variabilidad en la producción. Los municipios con  $R^2$  bajo ( $\leq 0.2$ ) son San Luis de La Paz (0.10), Santa Catarina (0.13), y Guanajuato (0.20) la superficie sembrada casi no explica la producción, lo que sugiere que otros factores dominan la productividad.

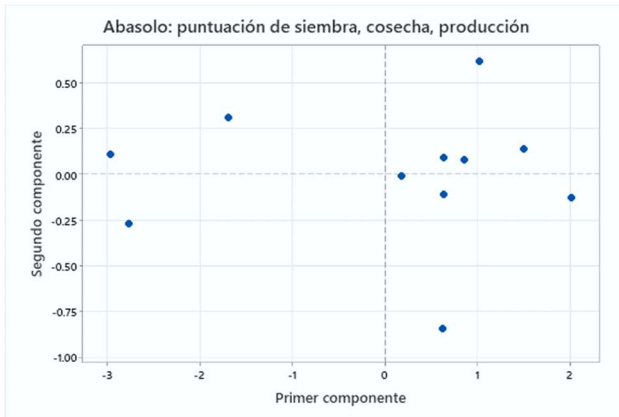


Figura 8. Gráfica de Abasolo.

Obtenida de: elaboración propia.

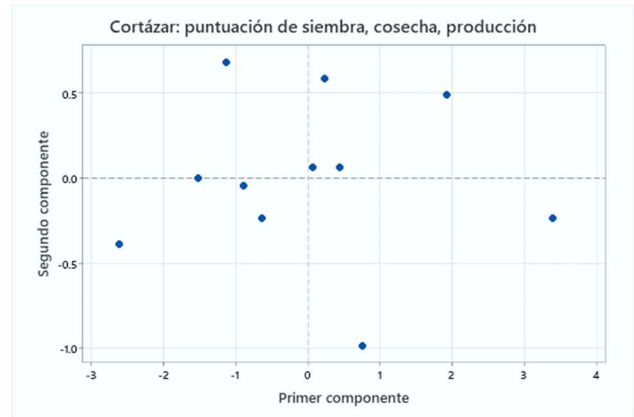


Figura 9. Gráfica de Cortázar.

Obtenida de: elaboración propia.

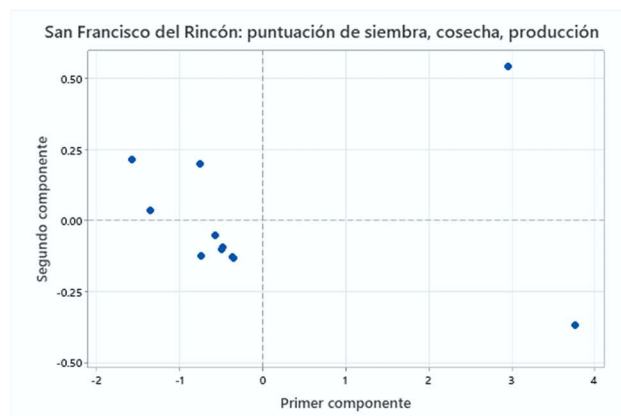


Figura 10. Gráfica de San Francisco del Rincón.

Obtenida de: elaboración propia.

### Análisis de Componentes Principales (PCA):

Se realizó un Análisis de Componentes Principales (PCA) obteniendo como resultados PC1 y PC2, que se explican a continuación. El componente principal 1 (PC1) explica la mayor

variabilidad observada (70%-95% en la mayoría de los municipios), asociado positivamente a: Área sembrada (Sembrada), Área cosechada (Cosechada) y Producción (toneladas). Indicando que la variabilidad en los datos se debe principalmente a la capacidad agrícola (siembra, cosecha y producción). La Tabla 1 muestra los resultados de Abasolo. La Tabla 2 muestra los resultados de Cortázar. La Tabla 3 muestra los resultados de San Francisco del Rincón. La Figura 8 muestra la gráfica PCA de Abasolo. La **¡Error! No se encuentra el origen de la referencia.** muestra la gráfica PCA de Cortázar. La **¡Error! No se encuentra el origen de la referencia.** muestra la gráfica PCA de San Francisco del Rincón.

Tabla 4. Resultados de Apaseo el Grande.

Apaseo el Grande				
Análisis de los valores de la matriz de correlación				
Valor propio	2.7428	1.0929	0.1644	0.0000
Proporción	0.686	0.273	0.041	0.000
Acumulada	0.686	0.959	1.000	1.000
Variable	PC1	PC2	PC3	PC4
Siembra	0.526	-0.459	0.266	0.665
Cosecha	0.592	-0.110	0.386	-0.699
Siniestro	-0.244	-0.865	-0.351	-0.265
Producción	0.560	0.171	-0.811	0.000

Obtenida de: elaboración propia.

Tabla 5. Resultados de Acámbaro.

Acámbaro				
Análisis de los valores de la matriz de correlación				
Valor propio	3.0112	0.8675	0.1213	-0.0000
Proporción	0.753	0.217	0.030	-0.000
Acumulada	0.753	0.970	1.000	1.000
Variable	PC1	PC2	PC3	PC4
Siembra	0.562	-0.181	0.403	0.699
Cosecha	0.568	-0.112	0.395	-0.714
Siniestro	-0.250	-0.967	-0.002	-0.048
Producción	0.547	-0.140	-0.825	0.000

Obtenida de: elaboración propia.

Tabla 6. Resultados de León.

León				
Análisis de los valores de la matriz de correlación				
Valor propio	2.1165	1.7166	0.1669	-0.0000
Proporción	0.529	0.429	0.042	-0.000
Acumulada	0.529	0.958	1.000	1.000
Variable	PC1	PC2	PC3	PC4
Siembra	-0.002	0.762	-0.148	-0.631
Cosecha	0.625	0.262	-0.582	0.450
Siniestro	-0.447	0.574	0.267	0.632
Producción	0.640	0.148	0.754	0.000

Obtenida de: elaboración propia.

El componente principal 2 (PC2) explica entre 4% y 30% de la variabilidad observada, dominado por: Área siniestrada (Siniestrada) con cargas positivas altas. Refleja las pérdidas por daños (plagas, falta de tecnología, etc.) son un factor secundario pero relevante en la variabilidad. La Tabla 4 muestra los resultados de Apaseo el Grande. La

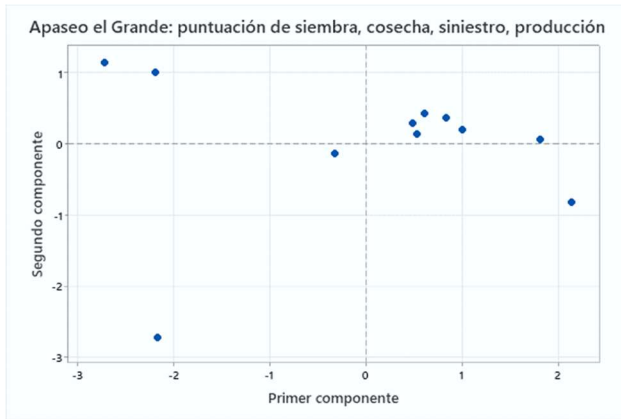


Figura 11. Gráfica de Apaseo el Grande.

Obtenida de: elaboración propia.

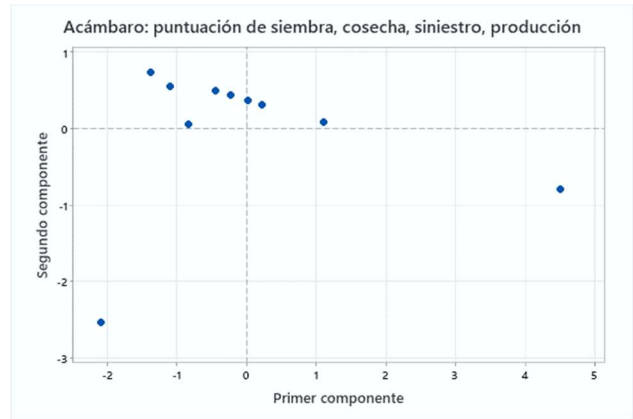


Figura 12. Acámbaro.

Obtenida de: elaboración propia.

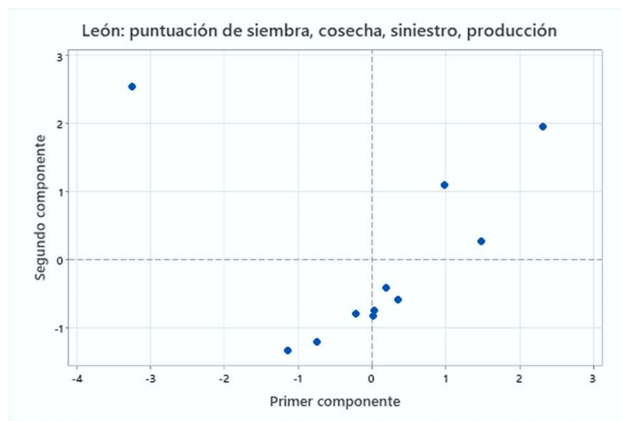


Figura 13. León.

Obtenida de: elaboración propia.

muestra los resultados de Acámbaro. La **¡Error! No se encuentra el origen de la referencia.** muestra los resultados de León. La Figura 11 muestra la grafica PCA de Apaseo el Grande. La Figura 12 muestra la grafica PCA de Acámbaro. La **¡Error! No se encuentra el origen de la referencia.** muestra la grafica PCA de León. El Siniestro tiene carga negativa en PC1, lo que sugiere que, a mayor producción, menor área dañada, o viceversa.

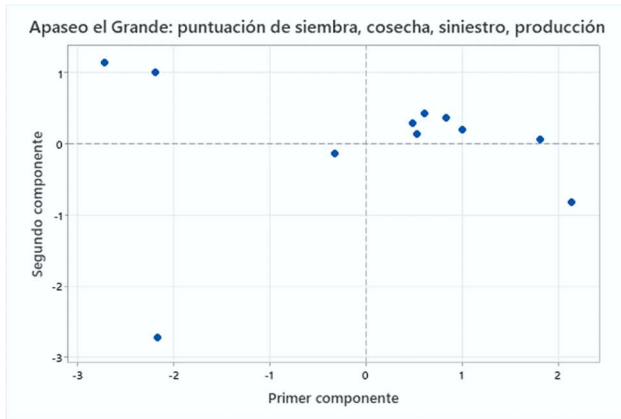


Figura 11. Gráfica de Apaseo el Grande.

*Obtenida de: elaboración propia.*

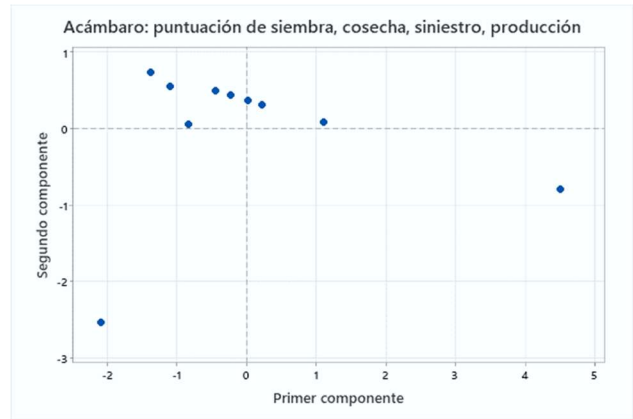


Figura 12. Acámbaro.

*Obtenida de: elaboración propia.*

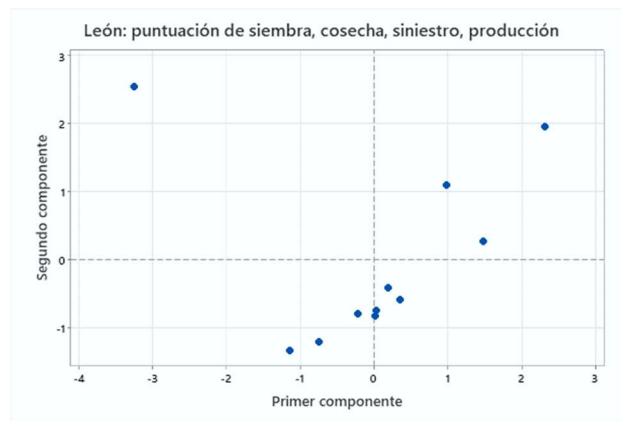


Figura 13. León.

*Obtenida de: elaboración propia.*

#### 4. Conclusiones: lo que podemos aprender de este artículo

El análisis multivariado aplicado a la producción de frijol reveló patrones críticos. La correlación superficie-producción mostró relaciones fuertemente positivas en municipios como Irapuato (0.986) y Villagrán (0.976), donde expandir el área cultivada impulsa la producción. En municipios como Victoria (-0.541) y Guanajuato (-0.452), esta relación fue negativa, sugiriendo limitaciones tecnológicas o degradación del suelo. La correlación superficie-rendimiento destacó ineficiencias: en Irapuato (-0.745) y Celaya (-0.614), mayor superficie redujo el rendimiento, probablemente por sobreexplotación o falta de tecnología. Los modelos de regresión confirmaron estas tendencias, con  $R^2$  alto en Irapuato (0.98) y bajo

en San Luis de La Paz (0.10), evidenciando que factores externos influyen donde la superficie no explica la producción. El PCA identificó que el 70-95% de la variabilidad obtenida (PC1) corresponde a variables agrícolas básicas (siembra, cosecha, producción), mientras que PC2 hasta 30% vinculó áreas siniestradas, resaltando el impacto de plagas y riesgos climáticos. La extrema desigualdad entre municipios como San Felipe vs Tierra Blanca refleja disparidades en acceso a tecnología o recursos. Estos hallazgos subrayan la urgencia de implementar tecnologías y automatización en las etapas del proceso agrícola para obtener un mayor rendimiento en producción. Además, se requiere capacitación en prácticas sostenibles, inversión en suelos marginales y políticas integradas que combatan las brechas socioeconómicas, asegurando así la seguridad alimentaria y sostenibilidad en la región.

## **5. Referencias: *por si quieres seguir conociendo más***

- Contreras-Medina, D. I., Contreras-Medina, L. M., Cerroblanco-Vázquez, V., Gallardo-Aguilar, M. C., González-Farías, J. P., Medina-Cuellar, S. E., Acosta-Montenegro, A., Lemus-Martínez, L. Y., Moreno-Ojeda, B., & Negrete-López, A. D. (2025). Reorienting Innovations for Sustainable Agriculture: A Study Based on Bean' s Traditional Knowledge Management. *Agriculture*, 15(560), 1–21. <https://doi.org/10.3390/agriculture15050560>
- Gallegos Acosta, J. A., Hernández C, Y., Jiménez, López Anaya, L. J., Montero, Victor, T. M., Brenda, C., & Guerrero Aguilar, Z. (2018). *Variedades De Frijol Y Producción De Semilla En Guanajuato*.
- Kadege, E. L., B Venkataramana, P., Assefa, T., Ndunguru, J. C., Rubyogo, J. C., & Mbega, E. R. (2024). Characterization of phenotypic traits associated with anthracnose resistance in selected common bean (*Phaseolus vulgaris* L.) breeding material. *Heliyon*, 10(5). <https://doi.org/10.1016/j.heliyon.2024.e26917>
- Pardo, C. E. (2020). *Estadística descriptiva multivariada* (A. M. Olaya Murillo, Ed.; Primera ed). Universidad Nacional de Colombia.
- Ping, C., Deng, X., Guo, Z., Luo, W., Li, X., & Xin, S. (2024). Characterizing the flavor profiles of Linjiangsi broad bean (*Vicia faba* L.) paste using bionic sensory and multivariate statistics analyses based on ripening time and fermentation environment. *Food Chemistry: X*, 23(July). <https://doi.org/10.1016/j.fochx.2024.101677>
- Rodriguez Espinoza, R. F., Ipanaqué Roña, J. M., Cruz Nieto, D. D., Abarca Rodriguez, J. J., Eguilas Caushi, Y. M., Gomero Mancesidor, J. M., Castro Bartolomé, H. J., Jamanca-Gonzales, N. C., & Silva-Paz, R. J. (2024). Influence of organic compost treatment on biometric patterns and sensory

- attributes of fresh green beans (*Phaseolus vulgaris* L.). *Case Studies in Chemical and Environmental Engineering*, 9(August 2023). <https://doi.org/10.1016/j.cscee.2024.100630>
- Servín-Palestina, M., Salazar-Moreno, R., López-Cruz, I., Medina-García, G., & Cid-Ríos, J. Á. (2022). Predicción de la producción y rendimiento de frijol, con modelos de redes neuronales artificiales y datos climáticos. *Biotecnia*, 24(2), 104–111. <https://doi.org/10.18633/biotecnia.v24i2.1664>
- Wiesinger, J. A., Osorno, J. M., McClean, P. E., Hart, J. J., & Glahn, R. P. (2021). Faster cooking times and improved iron bioavailability are associated with the down regulation of procyanidin synthesis in slow-darkening pinto beans (*Phaseolus vulgaris* L.). *Journal of Functional Foods*, 82(April), 104444. <https://doi.org/10.1016/j.jff.2021.104444>
- Zamukulu, P. M., Bagula, E. M., Mondo, J. M., Chuma, G. B., Safina, F. B., Cishesa, T. H., Kavange, A. B., Masumbuko, D. R., Kazadi, J. W., Mushagalusa, G. N., & Lubobo, A. K. (2023). Optimization of plant density and fertilizer application to improve biofortified common bean (*Phaseolus vulgaris* L.) yield on Nitisols of South-Kivu, Eastern D.R. Congo. *Heliyon*, 9(6). <https://doi.org/10.1016/j.heliyon.2023.e17293>