

ASOCIACION DE ALGUNAS CARACTERISTICAS CUANTITATIVAS CON EL RENDIMIENTO DE LA SOYA (*Glycine max*) EN EL TROPICO MEDIANTE DOS TECNICAS DE REGRESION 1/*

Enrique Villalobos *
Ramiro Barrantes **
Ronald Echandi ***

ABSTRACT

Association of some quantitative characters with soybean yield using two regression methods. The association of some quantitative characters with yield of soybean (*Glycine max* L. Merr) was investigated using two regression methods: path analysis coefficient and stepwise multiple regression. Field work was carried out at the Fabio Baudrit Experimental Field Station, in Alajuela, Costa Rica, at the end of the rainy season. Material was distributed in the field according to a random block design, with three replicates.

When the stepwise multiple regression was applied, it was concluded that the more determining variables on the yield of the soybean plant were the number of seeds per plant and the average seed weight. These variables amounted respectively 78% and 13% of total variation. With the path analysis coefficient technique, number of seeds per plant and the average seed weight showed to have the higher direct effect on plant yield (coefficients were 0.912 and 0.368 respectively) in agreement with the stepwise analysis. Using the latter technique, it was found that the days to bloom and the number of nodes and pods have a high positive indirect effect on plant yield through its influence on the number of seeds per plant, and in consequence, they can be used as visual indicators when selecting the most productive genotypes in tropical conditions.

INTRODUCCION

Las técnicas de regresión resultan de gran interés en el mejoramiento genético de las plantas si se usan para conocer las interrelaciones entre variables de importancia agronómica, especialmente de aquellas que más contribuyen con el rendimiento de las plantas. Una vez identificados los principales componentes del rendimiento, el proceso de selección visual de fenotipos más rendidores puede ser orientado en forma indirecta, por medio de esos

componentes. El rendimiento es un carácter cuantitativo de baja heredabilidad (2), de manera que la selección indirecta cobra especial interés si se fundamenta en aquellos componentes que están más estrechamente asociados con el mismo y que a su vez poseen una heredabilidad alta (5).

En 1971, Lal y Haque (11) emplearon por primera vez el método analítico de coeficientes de trayectoria ("path coefficient analysis"), para estudiar la asociación de algunas características de la planta con el rendimiento en la soya. A partir de esa fecha el empleo de esta técnica se ha generalizado en el estudio de características de importancia agronómica en esta leguminosa (3,9,14,15,16). La técnica de coeficiente de trayectoria resulta de gran utilidad ya que permite cuantificar los efectos directos e indirectos de un grupo de variables inde-

1/ Recibido para publicación el 20 de agosto de 1984.
* Centro de Investigaciones en Granos y Semillas (CIGRAS), Universidad de Costa Rica.
** Escuela de Biología, Universidad de Costa Rica.
*** Instituto Interamericano de Cooperación Agrícola (IICA), Costa Rica.

pendientes sobre una variable dependiente mediante la descomposición de los coeficientes de correlación. Para tal fin se hace necesario establecer un orden de causalidad entre las variables independientes ya sea *a priori* o con base en evidencia experimental (4,12,13).

El método de regresión múltiple escalonada ("Stepwise multiple regression") también se ha usado con el fin de analizar la contribución de algunas características morfológicas y fisiológicas con el rendimiento en las leguminosas de grano (1,8). A diferencia del método anterior este no requiere el establecimiento de una relación de causalidad entre las variables que se analizan. Por medio de este método la variable independiente que más influye sobre la variable dependiente es aquella que explica la mayor variabilidad de ésta y, en consecuencia, ocupará el primer lugar en la ecuación de regresión posteriormente cada lugar sucesivo será ocupado por la variable que determina la mayor proporción de la variabilidad de la variable dependiente que aún no es explicable por las variables que ya han sido incluidas en la ecuación (13). Adams (1) considera que esta técnica es muy versátil ya que permite hacer diferentes combinaciones de las variables bajo estudio con lo cual se pueden lograr las ecuaciones de regresión que más se ajustan a los objetivos del fitomejorador.

En el presente artículo se presenta información sobre el uso de las dos técnicas mencionadas en el estudio de la asociación de algunas variables cuantitativas con el rendimiento por planta de la soya (*Glycine max* L. Merr.) en condiciones tropicales.

MATERIALES Y METODOS

La parcela experimental para la realización del presente trabajo estuvo ubicada en la Estación Experimental Fabio Baudrit, en la provincia de Alajuela, a 10°01' de latitud norte y a 84°16' de longitud oeste, y a una altitud de 840 msnm. El trabajo tuvo lugar al final de la época lluviosa, septiembre a diciembre de 1980.

Los cultivares seleccionados para el trabajo fueron los siguientes:

Improved Pelican, Jupiter, Chung-Hsieng, Mandarin-com, Clark-63, Williams, Nanda, Negra-25 SPS, FAO-27396 y Brasil L-652. Algunos de los cultivares mencionados difieren en cuanto al lugar de origen y características agronómicas y representan los genotipos mejor adaptados disponibles a la fecha en Costa Rica.

Se usó un diseño experimental de bloques completos al azar, con 3 repeticiones, siguiendo la metodología empleada previamente en trabajos similares (9,11,16). Cada repetición estuvo constituida por 4 surcos de 3 m de largo, con una separación de 0,5 m entre ellos. La distancia entre plantas fue de 5 cm. La evaluación de las características se hizo con base en 12 plantas tomadas al azar de los 2 surcos centrales de cada parcela.

Las variables seleccionadas para asociar con el rendimiento se agrupan de la siguiente manera: 1) componentes básicos del rendimiento: número de vainas por planta, número de semillas por vaina, el peso promedio de la semilla y el número de semillas por planta 2) respuestas fotoperiódicas: número de días a la floración y a la madurez, altura de la planta, y número de nudos por planta y 3) otras: período de floración y altura del tercer nudo. Esta última se empleó para estimar la altura de las primeras vainas. Las características a evaluar se seleccionaron con base en varios trabajos de asociación de caracteres de la planta de soya con el rendimiento realizado en diferentes países (2,3,9,11,16).

Los análisis de regresión múltiple escalonada y de los coeficientes de trayectoria se hicieron siguiendo los programas del "Statistical Package for the Social Sciences" (13). Se hizo una transformación logarítmica de las variables: número de vainas por planta y número de semillas por planta, con el fin de normalizar su distribución.

El método de coeficientes de trayectoria se aplicó asociando únicamente cinco variables, las cuales fueron ordenadas en una secuencia de causalidad, con base en criterios puramente fisiológicos, con el fin de lograr la mayor objetividad posible. En la Figura 1 se presenta la relación de causalidad que se estableció. Los números del (1) al (5) indican el orden de causalidad, en donde cada variable es causa de la siguiente y a su vez efecto de la anterior. Las flechas dobles representan los coeficientes de correlación (r) entre las variables. Las flechas de un solo sentido corresponden a los coeficientes de trayectoria, y representan el efecto directo de cada una de las variables independientes sobre la variable dependiente. Estos coeficientes de trayectoria (P) son los coeficientes estandarizados de regresión (β) los cuales se obtienen directamente del análisis estadístico (13).

Los efectos indirectos de las diferentes variables sobre el rendimiento se calcularon siguiendo la metodología usada por Dewey y Lu (4), resol-

CUADRO 1. Análisis de regresión múltiple escalonada entre algunas características cuantitativas con el rendimiento por plantas de soya.

| Característica | R ² | Incremento en R ² | Valor de b | Error estándar de b | Valor de F | P |
|-------------------------------|----------------|------------------------------|------------|---------------------|------------|-------|
| Número de semillas por planta | 0,7815 | 78,15 | 18,99 | 0,53 | 1281,02 | 0,001 |
| Peso promedio de la semilla | 0,9141 | 13,26 | 46,27 | 1,97 | 550,47 | 0,001 |
| Número de nudos por planta | 0,9220 | 0,79 | 0,14 | 0,02 | 36,36 | 0,001 |
| Altura de la planta | 0,9244 | 0,24 | -0,01 | 0,01 | 11,09 | 0,001 |
| Altura del tercer nudo | 0,9250 | 0,06 | 0,05 | 0,03 | 2,76 | 0,025 |
| Número de vainas por planta | 0,9252 | 0,02 | -1,43 | 1,33 | 1,15 | 0,05 |
| Días a la floración | 0,9254 | 0,02 | 0,01 | 0,01 | 0,69 | 0,05 |
| Período de floración | 0,9261 | 0,07 | 0,07 | 0,02 | 3,27 | 0,005 |
| Días a la madurez | 0,9266 | 0,05 | -0,02 | 0,01 | 2,55 | 0,01 |
| Número de semillas por vaina | 0,9267 | 0,01 | -0,69 | 1,05 | 0,43 | 0,05 |

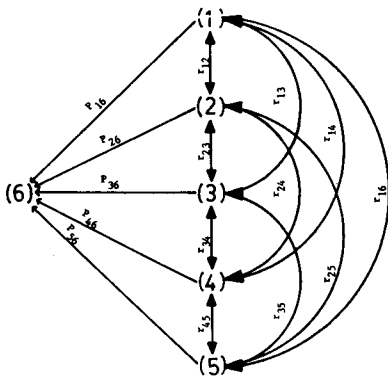


Fig. 1. Diagrama de trayectoria que muestra los coeficientes de correlación (flechas con doble cabeza) y los coeficientes estandarizados de regresión (flechas de una cabeza), mediante los cuales se obtuvieron los efectos directos e indirectos de cinco variables sobre el rendimiento de la soya. (1) Días a floración, (2) número de nudos por planta, (3) número de vainas por planta, (4) número de semillas por planta, (5) peso promedio de la semilla, (6) rendimiento por planta.

viendo las siguientes ecuaciones:

$$r_{16} = P_{16} + r_{12}P_{26} + r_{13}P_{36} + r_{14}P_{46} + r_{15}P_{56} \quad (1)$$

$$r_{26} = P_{26} + r_{12}P_{16} + r_{23}P_{36} + r_{24}P_{46} + r_{25}P_{56} \quad (2)$$

$$r_{36} = P_{36} + r_{13}P_{16} + r_{23}P_{26} + r_{34}P_{46} + r_{35}P_{56} \quad (3)$$

$$r_{46} = P_{46} + r_{14}P_{16} + r_{24}P_{26} + r_{34}P_{36} + r_{45}P_{56} \quad (4)$$

$$r_{56} = P_{56} + r_{15}P_{16} + r_{25}P_{26} + r_{35}P_{36} + r_{45}P_{46} \quad (5)$$

En todos los casos se hizo una observación visual de los residuos para determinar el ajuste de la regresión.

RESULTADOS Y DISCUSION

El ordenamiento de los resultados que se sigue en el Cuadro 1, obtenido mediante el análisis

CUADRO 2. Análisis de coeficientes de trayectoria que representa los efectos directos e indirectos de cinco variables cuantitativas sobre el rendimiento por plantas en soya*.

| Variable | Efecto por medio de: | | | | | |
|-----------------------------|----------------------|-------------------------|--------------------------|----------------------------|-----------------------------|---------------------------------------|
| | Días a la floración | No. de nudos por planta | No. de vainas por planta | No. de semillas por planta | Peso promedio de la semilla | Correlación con el rendimiento/planta |
| Días a la floración | <u>-0,014</u> | 0,075 | -0,040 | 0,344 | -0,024 | 0,341** |
| No. de nudos por planta | -0,006 | <u>0,160</u> | -0,071 | 0,710 | -0,050 | 0,743** |
| No. de vainas por planta | -0,006 | 0,129 | <u>-0,088</u> | 0,881 | -0,073 | 0,843** |
| No. de semilla por planta | -0,005 | 0,125 | -0,085 | <u>0,912</u> | -0,064 | 0,883** |
| Peso promedio de la semilla | 0,001 | -0,022 | 0,018 | -0,158 | <u>0,368</u> | 0,207** |

* Los valores subrayados corresponden a los efectos directos de cada variable sobre el rendimiento por planta.

de regresión múltiple escalonada, indica, en orden decreciente la contribución de cada una de las diez variables independientes estudiadas, al rendimiento por planta.

Este ordenamiento secuencial está dado por las modificaciones que sufre el coeficiente de determinación (R^2), conforme aumenta el número de variables en la ecuación de regresión. Puede observarse que el 92,67% de la variación del rendimiento por planta (variable dependiente) es atribuible a la influencia combinada de las diez variables bajo estudio. De éstas, sin embargo, las dos características más determinantes son: el número de semillas por planta, que explica 78,15% de la variación y el peso de la semilla (con 13,26%). Las ocho variables restantes sólo contribuyen con 1% de la variación del rendimiento por planta.

Por medio del análisis de regresión de trayectoria, Figura 2, se llegó a determinar que el número de semillas por planta es el componente del rendimiento que mostró el mayor efecto directo sobre el rendimiento por planta (0,912) seguido del peso promedio de la semilla (0,368), según se consigna en el Cuadro 2. Estos resultados corroboran los obtenidos por medio de la regresión múltiple escalonada. Se concluye entonces que aquellas características que muestran el mayor efecto directo sobre el rendimiento son, a su vez, las que causan la mayor parte de la variabilidad en la variable dependiente. Los resultados obtenidos en el presente trabajo coinciden con investigaciones similares conducidas por tres años consecutivos en Suiza (10).

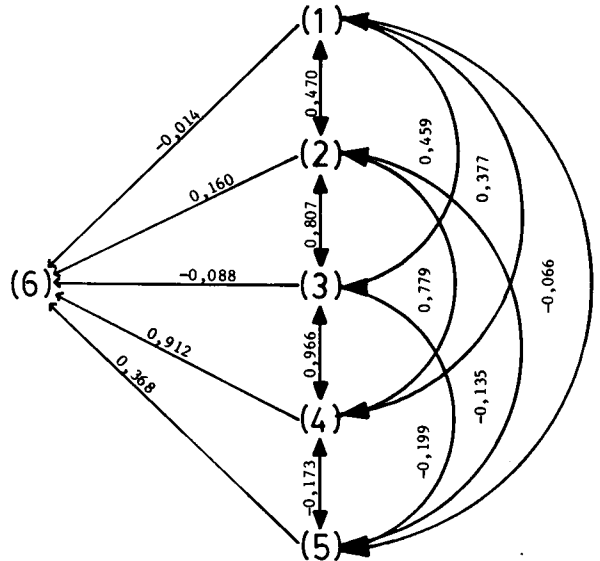


Fig. 2. Diagrama de trayectoria con los valores de los coeficientes de trayectoria y de correlación de cinco variables que afectan al rendimiento por planta en soya. (1) Días a la floración, (2) número de nudos por planta, (3) número de vainas por planta, (4) número de semillas por planta, (5) peso promedio de la semilla, (6) rendimiento por planta.

Resulta lógico, por lo tanto, que sean el número de semillas por planta y el peso promedio de la semilla, las dos variables que más contribuyen al rendimiento por planta, puesto que el producto de ambos constituye el rendimiento propiamente. Ese hecho podría conducir a menospreciar como elementos de selección visual algunas características de interés agronómico que podrían ser de gran im-

portancia por sus efectos determinantes del rendimiento, mediante la modificación del número o el peso de las semillas. Desde este punto de vista, el análisis de coeficientes de trayectoria resultó de suma utilidad, pues se puso en evidencia que otras características como el número de días a la floración, el número de nudos por planta y el número de vainas por planta presentan efectos indirectos altos sobre el rendimiento por planta, por su incidencia en el número de semillas por planta (Cuadro 2); por ello pueden ser de gran utilidad en la selección visual de genotipos de mayor rendimiento en la soya. Este aspecto cobra más importancia si además se toma en cuenta que los días a la floración y el número de nudos por planta tienen una heredabilidad más alta que el rendimiento por planta (6,7).

El efecto positivo del peso de la semilla sobre el rendimiento se vio parcialmente anulado por el número de semillas por planta (efecto indirecto -0,158), según se ilustra en el Cuadro 2. Consecuentemente, el efecto directo del peso de la semilla sobre el rendimiento (0,368) estuvo contrarrestado y el coeficiente de correlación entre el peso de la semilla y el rendimiento, aunque significativo, resultó relativamente bajo (0,207).

Se deduce de la presente investigación que la técnica de análisis de coeficientes de trayectoria es de suma utilidad ya que permite conocer la contribución real de un grupo de variables independientes sobre una variable dependiente, por medio de sus efectos directos e indirectos. Esta cualidad propia de esta técnica cobra mayor significado si se usan pocas variables y se ordenan siguiendo criterios fisiológicos, como se hizo en el presente estudio.

El establecimiento de un diagrama de causalidad con un número grande de variables es difícil de establecer con objetividad. Este aspecto ha sido analizado ampliamente por Li (12). Por su parte, el método de regresión múltiple escalonada no requiere un ordenamiento de causa y efecto entre las variables que se estudian, lo cual permite comparar simultáneamente un mayor número de características; sin embargo, no es posible, por medio de este método, conocer si una variable influye positivamente sobre la variable dependiente, a través de la modificación de otras características que se incluyen en la ecuación.

En el presente estudio ambas técnicas resultaron complementarias, aumentando la confiabilidad en las conclusiones finales.

RESUMEN

Se estudió la asociación de algunas características cuantitativas con el rendimiento por planta de soya (*Glycine max* L. Merr.), mediante dos métodos de regresión: análisis de coeficientes de trayectoria y regresión múltiple escalonada. El trabajo de campo se realizó en la Estación Experimental Fabio Baudrit en Alajuela al final de la época lluviosa. Los materiales se distribuyeron en el campo de acuerdo a un diseño experimental de bloques completos al azar con tres repeticiones.

Aplicando la técnica de regresión múltiple escalonada se logró determinar que las dos variables más determinantes del rendimiento fueron el número de semillas por planta y el peso promedio de la semilla, las que contribuyeron con el 78% y el 13% de la variación. Con la técnica de coeficientes de trayectoria se logró comprobar que el número de semillas por planta y el peso promedio de la semilla mostraron los efectos directos más altos sobre el rendimiento por planta (coeficientes de 0,912 y 0,368 respectivamente) en apoyo de los resultados obtenidos con el análisis de regresión múltiple escalonada. Mediante esta última técnica se logró establecer que los días a la floración y el número de nudos y de vainas tienen un efecto positivo indirecto alto sobre el rendimiento, al afectar el número de semillas por planta. Por lo tanto, pueden servir como indicadores visuales para la selección de genotipos más productivos en condiciones tropicales.

LITERATURA CITADA

1. ADAMS, M.W. y REICOSKY, D.A. Plant architecture and physiological efficiency in the field bean. Report to the Rockefeller Foundation. Michigan State University. Department of Crop Science, pp. 61-90, 1975.
2. BRIM, C.A. Quantitative genetics and breeding. *In* Soybeans: Improvement, Production and Uses. B.E. Caldwell, ed. Madison, Wisconsin, American Society of Agronomy, 1973. pp. 155-186.
3. CHAUDHARY, D.N. y SINGH, B.B. Correlation and path coefficient analysis of yield components in soybean. *Indian Journal of Agricultural Science* 44: 487-490. 1974.
4. DEWEY, D.R. y LU, K.H. A correlation and path coefficient analysis of components of crested wheat grass seed production. *Agronomy Journal* 51:515-518. 1959.

5. FALCONER, D.S. Introducción a la genética cuantitativa. México, Continental, 1970. 430 p.
6. JOHNSON, H.W., ROBINSON, H.F. y COMSTOCK, R.E. Estimates of genetic and environmental variability in soybeans. *Agronomy Journal* 47: 314-318. 1955.
7. JOHNSON, H.W. y BERNARD, R.L. Soybean genetics and breeding. In *The Soybean*. A.G. Norman, ed. New York, Academic Press, 1963. pp. 1-73.
8. JOSHI, J.M. y SMITH, P.E. Correlated response of certain plants traits with yield in soybeans. *Soybean Genetics Newsletter*. 5: 62-66. 1978.
9. KAW, R.N. y MENON, P.M. Association between yield and components in soybean. *Indian Journal of Genetics and Plant Breeding* 32: 276-280. 1972.
10. KELLER, E.R., SOLDATI, A. y PIATTINI, E. Study on the technique of crossing as well as on the genetics behaviour of quantitative characters of soybeans. *Soybean Genetics Newsletter*. 5: 79-80. 1978.
11. LAL, V.S. y HAQUE, M.F. Path analysis of yield components in soybeans. *Indian Journal of Genetics and Plant Breeding* 31: 357-362. 1971.
12. LI, C.C. Path analysis: A primer. California, Boxwood Press, 1977. 347 p.
13. NIE, N.H., HULL, C.H., JENKINS, J. G., STEINBRENNER, K. y BENT, D.H. Statistical package for the Social Sciences. New York, Mc Graw-Hill, 1975. pp. 320-397.
14. PANDEY, J.P. y TORRIE, J.H. Path coefficient analysis of seed yield components in soybeans (*Glycine max* (L) Merr.). *Crop Science* 13: 505-507. 1973.
15. PATIL, V.N. y POKLE, Y.S. Path coefficient analysis in soybean. *Indian Journal of Agricultural Science* 44: 544-546. 1974.
16. SENGUPTA, K. y KATARIA, A.S. Path coefficients analysis for some characters in soybeans. *Indian Journal of Genetics and Plant Breeding* 31: 292-297. 1971.