Aplicación del Análisis de Senderos en Estudios Biológicos

Adalberto Benavides-Mendoza*, Homero Ramírez, Valentín Robledo-Torres, José Hernández-Dávila

Departamento de Horticultura, Universidad Autónoma Agraria Antonio Narro, Buenavista, 25315, Saltillo, Coah., México. E-mail: abenmen@uaaan.mx (*Autor responsable).

Abstract

A revision of the theory of the paths analysis method is here presented, as a subsequent analysis of the correlation matrix, when a causal structural model is utilized. The differences with the non-causal regression models are discussed, as well as the advantages, and disadvantages, of path analysis. Some applications in ecological, agronomical, and plant breeding studies are revised.

Keywords: Causal models, correlation, regression, beta coefficients, path coefficient.

Resumen

Se presenta una revisión de la teoría del método de análisis de senderos, como un análisis posterior de la matriz de correlaciones, cuando se utiliza un enfoque causal de un modelo estructural o de factores. Se discuten las diferencias con los modelos no causales de regresión así como las ventajas y desventajas del método de senderos. Se revisan algunas aplicaciones en estudios ecológicos, agronómicos y de mejoramiento de plantas.

Palabras clave. Modelos causales, correlación, regresión, coeficientes beta, coeficientes de sendero.

Introducción

La construcción de matrices de correlación es común en estudios de sistemas biológicos. El análisis posterior de la matriz de correlaciones, utilizando modelos estructurales o de factores, permite profundizar en la clase de relaciones que se presentan entre las variables consideradas y en los efectos resultantes de dichas relaciones. El objetivo del presente trabajo es presentar una revisión de la teoría del método de análisis de senderos así como de algunas de las aplicaciones de dicha técnica de análisis en los estudios biológicos.

El análisis de senderos (*path analysis*) es una clase de modelo causal que combina las características de los modelos estructurales y de los de regresión múltiple. El análisis de senderos se utiliza ampliamente en las ciencias sociales, si bien la aplicación en las ciencias biológicas se ha extendido fuertemente, sobre todo en las áreas de Genética, Ecología y Fisiología.

Este trabajo tiene como objetivo revisar las bases teóricas que definen la extensión de aplicación del análisis de senderos y presenta algunos ejemplos de utilización en estudios de sistemas biológicos.

La Matriz de Varianza-covarianza y la Matriz de Correlaciones

Si partimos de una base de datos multivariada con *k* variables y *n* observaciones por variable, entonces tenemos una matriz de datos X de orden (n x k). Cuando esta matriz X contiene datos corregidos respecto a la media, esto es $(x - \bar{x})$ entonces la siguiente expresión es el estimador de la matriz de varianza-covarianza de orden (k x k):

$$S = (n-1)^{-1} X'X \qquad \text{(ec. 1)}$$

Por otro lado, si la matriz X contiene como elementos datos expresados en forma estandarizada, es decir $s^{-1}(x -)$, entonces el producto siguiente:

$$R = (n-1)^{-1} X'X \qquad \text{(ec. 2)}$$

Es el estimador de la matriz de correlaciones de orden (k x k).

Los elementos de la matriz R son los coeficientes de correlación lineal simple entre cada par de variables. El coeficiente de correlación describe, en forma simétrica

(es decir, respecto a ambas variables) la asociación lineal presente entre las variables. Mientras mayor es la magnitud absoluta del coeficiente de correlación más grande será la asociación entre las variables (Li, 1975; Kendall y Stuart, 1976; Pindyck y Rubinfeld, 1981).

Los Modelos de Regresión Múltiple

Los modelos de regresión lineal múltiple son modelos lineales utilizados para describir matemáticamente la relación entre una variable **Y**, llamada variable de respuesta, y una serie de variables predictoras, $\mathbf{X_1}$, $\mathbf{X_2}$, ..., $\mathbf{X_k}$, cuyos valores pueden ser fijos o aleatorios. Se supone que el valor específico de la variable de respuesta puede predecirse con cierta aproximación a partir de los valores que toman las variables predictoras. En otras palabras, el modelo de regresión sirve como resumen de la variación observada en la variable de respuesta en función de las variables predictoras (Kendall y Stuart, 1976; Draper y Smith, 1981; Pedhazur, 1982).

Un modelo lineal como el siguiente:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_k X_k \qquad \text{(ec. 3)}$$

dice que el valor de la variable de respuesta es una función lineal de las variables predictoras. Los coeficientes β_j, j=1,...k, son los llamados coeficientes de regresión parcial e indican la cantidad de cambio esperado en la variable de respuesta, Y, por unidad de cambio en la j-ésima variable predictora. Desde el punto de vista estadístico, cumpliendo ciertas suposiciones, el producto $\beta_j X_j$, j=1,...,k, es el mejor estimador del valor esperado de Y en función del valor de la variable X_j, manteniendo fijas las restantes variables predictoras (Draper y Smith, 1981). Cuando el modelo de la ecuación 3 utiliza variables estandarizadas entonces los coeficientes β_j, j=1,...k, son los llamados coeficientes beta. La interpretación es la misma marcada para los coeficientes de regresión parcial excepto que las unidades son de desviación estándar. Estos coeficientes beta no son coeficientes de sendero a no ser que se asocien a un modelo estructural con vías causales.

Al contrario de lo que ocurre con el análisis de correlación en el de regresión la relación descrita no es simétrica, ya que implica un sentido de las variables predictoras hacia la de respuesta, expresando esta última como una función lineal de aquellas. Los valores âj son obtenidos utilizando una técnica especial, llamada de cuadrados mínimos, con la cual se minimizan las diferencias entre los valores de Y realmente observados y los valores generados por una función lineal como la que aparece en la ecuación 3.

En su forma tradicional el análisis de regresión en ningún momento pretende ser un modelo funcional de la situación que describe; simplemente puede ser un buen modelo en el sentido de minimizar los errores de predicción de Y con base en la información dada por las variables predictoras. Dicho de otra forma, *la orientación básica de los modelos de regresión es la predicción*. La extensión de su uso hacia la construcción de modelos funcionales o estructurales de la situación bajo análisis se debe realizar con las necesarias precauciones (Li, 1975; Draper y Smith, 1981; Chanter, 1981; Pindyck y Rubinfeld, 1981).

El funcionamiento del modelo clásico de regresión puede resumirse con el esquema que aparece en la Figura 1:

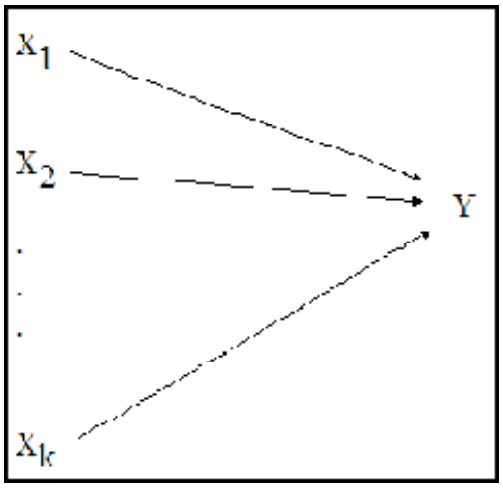

Figura 1. Esquema del modelo clásico de regresión lineal múltiple.

En donde se representa una variable aleatoria Y, cuyo valor depende o está condicionado por los valores tomados por las variables X_1, X_2,..., X_k. En el modelo clásico de regresión las variables predictoras son de valor fijo y no están correlacionadas entre sí (Draper y Smith, 1981; Chanter, 1981). El ejemplo más conocido de una situación lo constituye un diseño experimental con arreglo ortogonal (como un arreglo factorial). Una suposición muy importante en el modelo clásico de regresión, suposición generalmente no observada en estudios de sistemas bajo condiciones no controladas, es la ausencia de correlación entre las variables predictoras. Las propiedades óptimas de estimación basadas en el ajuste de cuadrados mínimos se obtienen bajo la situación de ortogonalidad (Hoerl y Kennard, 1970; Draper y Smith, 1981; Chanter, 1981; Stewart, 1987).

Análisis de Senderos

El método de análisis de senderos, llamado también método de senderos, o método de coeficientes de sendero, o de vías causales fue desarrollado originalmente por Wright (1921) para el análisis de sistemas cerrados de

variables unidas por relaciones causales. Específicamente, Wright (1921, 1934) desarrolló esta técnica como apoyo en la interpretación de los datos provenientes de sistemas que, por una u otra razón, eran difíciles de manipular experimentalmente.

Si bien el desarrollo principal del análisis de senderos corresponde a Wright (1921), aproximadamente en la misma época y de manera independiente, Ronald Aymer Fisher había desarrollado el concepto de factores de un coeficiente de correlación, concepto idéntico al de los coeficientes de sendero de Wright (Li, 1968).
Diferentes autores describieron y desarrollaron la teoría subyacente a dicho análisis y el lector interesado en los aspectos matemáticos formales es remitido a los trabajos de Kempthorne (1969), Li (1956, 1975), Turner y Stevens (1959) y Wright (1921, 1934, 1960). Otros trabajos con nivel matemático más accesible y con ejemplos específicos de aplicación son los de Asher (1983), Crespi (1990), Pedhazur (1982), Reyes (1992) y Reyes y Benavides (1993).

En su forma original, tal como es descrito por Wright (1921, 1934) la técnica de análisis de senderos tiene una orientación eminentemente biológica; de hecho el propósito primario del método fue eliminar hasta cierto punto los problemas surgidos al no cumplir los supuestos de los métodos estadísticos, como la regresión múltiple, en los estudios biológicos. Crespi (1990) realizó una excelente recapitulación de las diferencias entre los modelos causales con orientación puramente estadística (predictivos) y los que tienen orientación biológica (explicativos).

El análisis de senderos es una forma de análisis de regresión lineal estructurado con respecto a variables estandarizadas en un sistema cerrado formalmente completo (Wright, 1921; Li, 1975). El método de senderos no concierne tanto con la predicción como con la *proposición* de una interpretación plausible de las relaciones entre las variables estudiadas. En otras palabras, el análisis de senderos concierne con la construcción de una estructura causal compatible con los datos observados (Wright, 1921; Li, 1975; Pedhazur, 1982).

Lo anterior marca los límites de aplicación del análisis de senderos. Este no debe de aplicarse para construir modelos de relaciones entre variables a partir de series de datos, más bien es lo contrario, el análisis es aplicable *a posteriori* de la propuesta de una estructura causal entre variables para investigar la importancia de efectos causales directos, e indirecto, en un sistema de variables correlacionadas (Wright, 1921, 1934; Li, 1975; Crespi, 1990). Mitchell (1992, p. 123) comenta: "Briefly, path analysis allows one to partition the simple correlations among a set of variables according to a particular working model about their causal relationships..."

Se hace obvio entonces el principal inconveniente del análisis de senderos: los resultados serán tan buenos como adecuado sea el modelo causal propuesto para el sistema bajo estudio. El uso de modelos inadecuados nos llevará a conclusiones erróneas (MacDonald, 1977; Mitchell, 1992; Dofing y Knight, 1992). Los errores no se referirán, en este caso, a las estimaciones de los coeficientes de un modelo, más bien se relacionarán con la lógica de la situación bajo análisis (véase como ejemplo el artículo de Dofing y Knight, 1992). El problema en este caso es que no se ha desarrollado un método sencillo para la comparación clara de diferentes modelos de un mismo sistema (llamados modelos en competencia) ya que, hasta ahora, no existe una forma directa y simple de cuantificar el acuerdo entre el diagrama causal (el modelo) y los datos sobre los que se trabaja (Pedhazur, 1982; Mitchell, 1992).

Kempthorne (1969), en el área de la genética, que fue en donde primeramente se aplicó el análisis de senderos, no recomendó el uso extensivo de esta clase de análisis dada la excesiva facilidad con que genera soluciones atractivas a partir de modelos erróneos. Para una interesante exposición acerca del uso indiscriminado del análisis de senderos en el área de las ciencias sociales véase Pedhazur (1982).

Como una manera de resolver la carencia de una prueba objetiva con la cual verificar el ajuste de la estructura causal propuesta a la estructura de covarianzas observada, Mitchell (1992) propuso extender la solución del análisis de senderos hacia la aplicación de modelos estructurales lineales LISREL. En un modelo LISREL la solución para los coeficientes de sendero se obtiene utilizando métodos de máxima verosimilitud, cuadrados mínimos u otros. Las soluciones se obtienen a partir de un conjunto de restricciones de identificación (representadas por el diagrama causal bajo consideración) basadas en la estructura de covarianzas del problema que se plantea. La gran ventaja de un modelo LISREL es que se dispone de una forma directa de verificar la bondad del ajuste de cada modelo en competencia respecto a la estructura observada de covarianzas; su principal desventaja es que son bastante más complicados si se comparan con la sencillez de la solución clásica de los coeficientes de sendero.

En cuanto a las suposiciones sobre las que se basa el análisis de senderos, éstas son expuestas de manera excelente en los trabajos de Wright (1921, 1934), Li (1956, 1975), Pedhazur (1982) y Asher (1983). Que tanto estas suposiciones son consideradas antes de realizar un análisis de senderos sobre un grupo de datos? La situación es la misma que se presenta para la regresión múltiple: buena

parte de los usuarios o bien no las consideran o las atienden solo en parte (vease nota de Kang, 1992).

Estructuras Causales

El modelo causal propuesto de acuerdo a cierta teoría de las relaciones entre las variables se puede trasladar de manera sucinta a un diagrama causal en donde, por medio de flechas indicativas uni- o bidireccionales, se indican las supuestas relaciones causales presentes entre las variables. El diagrama causal es traducido en forma matemática construyendo un sistema de ecuaciones lineales (que llevan o no restricciones) que refleja la estructura propuesta.

Considérese el diagrama causal de la Figura 2 que muestra un modelo de un sistema hipotético con relaciones causales:

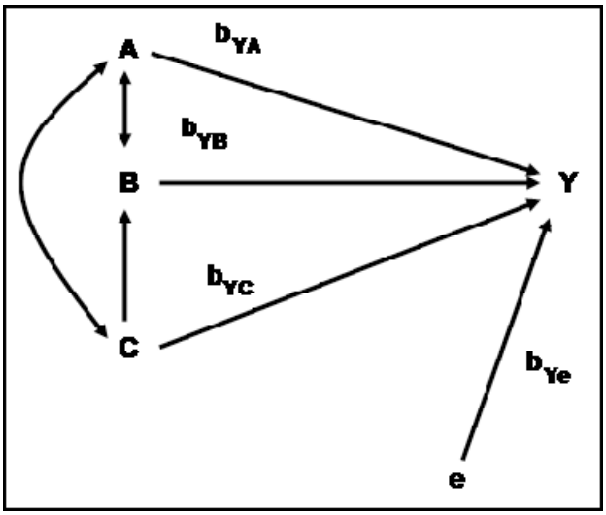

Figura 2. Diagrama causal de un sistema de variables hipotético.

Dicho diagrama muestra un arreglo de interacciones entre variables. La variable de respuesta, Y, se encuentra determinada por las variables predictoras (llamadas también variables básicas) A, B, C y por un efecto residual, e, que agrupa a todos aquellos factores diferentes a A, B y C y que tengan alguna influencia sobre Y. Las variables A, B y C se encuentran correlacionadas entre sí y con la variable Y. Por definición, el término residual no se correlaciona con las variables predictoras. La inclusión del término residual es el artificio con el cual el sistema se convierte en uno formalmente cerrado (Wright, 1921; Turner y Stevens, 1959; Kempthorne, 1969; Li, 1956, 1971; Pedhazur, 1982). El modelo presentado es de tipo recursivo es decir, el flujo causal propuesto para el sistema es unidireccional y queda fuera de consideración la causalidad recíproca. Se recomienda el texto de Asher (1983) para revisar su exposición acerca de modelos no recursivos.

Las variables predictoras y las de respuesta se encuentran enlazadas en una estructura causal correlativa que se supone es descrita por el modelo causal propuesto. El análisis de senderos permite fraccionar en factores la correlación observada entre las variables predictoras y la de respuesta obteniendo del valor total una serie de componentes que son, para el modelo analizado (Li, 1975; Pedhazur, 1982):

a). Un efecto causal directo: b_{YA}, b_{YB} ó b_{YC}.
b). Un efecto causal indirecto: b_{CB}.
c). componentes no causales que son no analizables o espurias.
d). un componente residual: b_{YE}.

La partición del valor del coeficiente de correlación no implica que necesariamente aparezcan todas estas componentes, ya que su presencia o ausencia depende enteramente de la estructura causal especificada (Mitchell, 1992).

A partir del modelo causal anterior la partición de los coeficientes de correlación es descrita por el siguiente sistema de ecuaciones lineales:

$$
\begin{aligned}
r_{AY} &= \underset{\text{directo}}{b_{YA}} + \underset{\text{no analizable}}{r_{AB}b_{YB}} + \underset{\text{no analizable}}{r_{AC}b_{YC}} \\
r_{BY} &= \underset{\text{no analizable}}{r_{BA}b_{YA}} + \underset{\text{directo}}{b_{YB}} + \underset{\text{espurio}}{r_{CB}b_{YC}} \qquad \text{(ec. 4)} \\
r_{CY} &= \underset{\text{no analizable}}{r_{AC}b_{YA}} + \underset{\text{indirecto}}{r_{BC}b_{YB}} + \underset{\text{directo}}{b_{YC}}
\end{aligned}
$$

La descomposición de los coeficientes de correlación genera una serie de componentes directos b_{YA}, b_{YB} y b_{YC}, cuyo valor numérico representa la magnitud o importancia del efecto causal directo de cada variable predictora sobre la variable de respuesta. A este componente directo se le llama coeficiente de sendero de la vía causal referida, que puede ser, como ejemplo, la vía $A \rightarrow Y$.

Se tienen posteriormente una serie de componentes bajo el rubro de no analizables. En este caso lo que ocurre es que se carece de la información suficiente a partir de la cual realizar, con la certidumbre necesaria, una posterior descomposición de estos componentes; es por ello que, a estos últimos, les asignamos únicamente magnitud y no una dirección causal precisa, tal como ocurre en el caso de la vía $A \leftrightarrow B$.

El componente indirecto surge como consecuencia de la acción de dos variables (una siendo causa y la otra efecto) sobre una tercera que resulta de la acción de las dos primeras. En nuestro ejemplo la magnitud y dirección causal representan el efecto que ejerce la variable C sobre la variable de respuesta, Y, a través de la variable predictora

B. En otras palabras, el producto $r_{BC}b_{YB}$ representa la parte de la correlación entre las variables C y Y que resulta de la acción de C sobre B.

El componente que llamamos espurio representa la parte "falsa" de la correlación entre dos variables que resulta del hecho de que ambas tienen una causa común. En nuestro ejemplo el componente espurio es resultado de que las variables B y Y tienen una causa común que es la variable C.

Por último, el efecto del componente residual se calcula por diferencia entre la unidad y el producto escalar $\mathbf{b'R}_y$, operación que es equivalente a restarle a la unidad el coeficiente de determinación ($1-R^2$). Como fue explicado, este componente residual es el artificio matemático que permite que nuestro sistema causal sea un sistema cerrado.

Aplicaciones en Estudios Biológicos

A continuación aparecen, agrupados en áreas de estudio afines, una serie de trabajos en donde se aplicó el análisis de senderos. Para cada trabajo, ya sea de manera individual o para un conjunto de ellos, se anota una breve descripción del mismo. Los trabajos citados se refieren a las aplicaciones en ciencias biológicas como la Ecología, Fisiología, Agronomía, etc.

Aplicaciones en Fisiología y Ecología Vegetal

Farris y Lechowicz (1990). En este trabajo se relacionan una serie de caracteres morfológicos, fenológicos y fisiológicos con la capacidad reproductiva de *Xanthium strumarium* (Compositae). Se verificaron las vías causales que confluyen hacia la diferenciación y crecimiento de las estructuras reproductivas.

Leduc *et al.* (1992). Se estudió la distribución de especies en una comunidad forestal. Los autores proponen una metodología que, haciendo uso de análisis de senderos y pruebas parciales de Mantel, permite separar, desde el punto de vista matemático, la componente puramente espacial de aquella otra determinada por las condiciones ambientales como factores causales de la distribución observada de individuos de especies arbóreas.

Mitchell-Olds (1987). Aplicación del análisis de senderos en el estudio de la competencia intraespecífica. Se describen las ventajas del método de senderos sobre otra serie de técnicas de análisis de los datos. Es un reporte de calidad y se recomienda para consultar la aplicación del método de senderos en estudios ecológicos.

Schemske y Horvitz (1988). Trabajo en donde se describieron las interacciones entre ciertas especies de insectos y una especie vegetal, *Calathea ovandensis* (Marantaceae). De acuerdo a observaciones previas se conocía cierto efecto de los insectos sobre la producción de flores y frutos. Se construyó un diagrama causal basado en esas observaciones y se analizaron una serie de datos utilizando análisis de senderos. Es un reporte cuya lectura es recomendable.

Sun (1989). Estudio acerca de la relación entre los niveles endógenos de poliaminas y el número y tasa de iniciación de raíces laterales en plántulas de manzano.

Wesser y Armbruster (1991). Un trabajo excelente acerca de la prueba de dos hipótesis en competencia respecto a la distribución observada de especies vegetales en comunidades específicas. Se aplicaron modelos causales y experimentación directa para verificar las consecuencias derivadas de cada modelo. Es un ejemplo del tratamiento de modelos en competencia.

Abrol (1990). Estudio del efecto de diferentes condiciones ambientales sobre la actividad de polinización por himenópteros en plantas de alfalfa.

Pantone *et al.* (1989). Se estudió la eficacia del nemátodo *Anguina ansinckiae* como agente para el biocontrol de la maleza *Amsinckia intermedia* en un experimento de competencia con plantas de trigo.

Shasha'a *et al.* (1973). Investigación sobre la relación entre la producción de semilla de *Allium cepa* L. y su interrelación con la actividad de insectos polinizadores y los componentes del rendimiento del cultivo.

Talukdar *et al.* (1991). Estudio de la dinámica poblacional de chicharritas (*Nephottettix* sp.) y el efecto que tiene sobre este proceso la aplicación de fertilizantes específicos a las plantas hospederas. Se estudiaron también las correlaciones entre los niveles de población de los depredadores y los de *Nephotettix* sp. para obtener, aplicando un análisis de senderos, el impacto relativo de cada tipo de depredador.

Aplicaciones en Selección Natural y Genética Poblacional

Crespi (1990). Trabajo acerca de la acción de la selección natural sobre sistemas genéticos correlacionados. Este artículo es un trabajo magnífico sobre modelaje matemático de la selección natural y su lectura es muy recomendable.

Jordan (1992). Se realizó un estudio acerca de la adaptación de ecotipos morfológicamente distintos de *Diodia teres* (Rubiaceae). Se aplicó un análisis de senderos a las observaciones considerando un diagrama causal en donde una serie de caracteres tienen influencia sobre la aptitud biológica de los individuos.

Mitchell-Olds y Bergelson (1990). Estudio sobre la variabilidad de una serie de caracteres cuantitativos en una población de *Impatiens capensis*. El análisis de senderos se aplicó para investigar los efectos directos y

los resultantes de las correlaciones de los caracteres sobre la aptitud biológica de los individuos. Este artículo es un trabajo notable de genética estadística aplicada a estudios evolutivos y su lectura es muy recomendable.

Pritts y Hancock (1985). Trabajo cuyo objetivo fue examinar la dinámica de la reproducción de una especie vegetal perenne. Se estudió la relación entre el reparto de fotosintatos a diferentes sitios de utilización reproductivos y la edad de los individuos.

Wright (1921); Li (1956). Aplicación de los coeficientes de sendero en estudios de genética poblacional. El trabajo de Li (1956) incluye datos históricos acerca del desarrollo teórico del análisis de senderos. El trabajo de Wright (1921) es el artículo original en donde se establecieron las bases del método. Estos reportes son de lectura obligada para los genetistas de poblaciones y mejoradores de plantas.

Aplicaciones en Nutrición Vegetal y Fertilidad de Suelos

Ferrari (1963). Descripción y comparación del análisis de senderos frente al modelo convencional de regresión múltiple. Se utilizó como ejemplo de aplicación un modelo causal que involucra variables que son causa y a su vez efecto de otras variables. Los datos utilizados provienen de una investigación realizada acerca del efecto de la fertilidad del suelo y otros factores sobre el contenido de Mg y K de plantas forrajeras. El estudio descrito es una investigación de campo con factores no controlados; su lectura es recomendable.

May *et al.* (1991). Estudio en relación a los diferentes factores fisiológicos y ambientales que influyen sobre la eficiencia de absorción y reparto interno de nitrógeno en la planta de trigo.

Prasad *et al.* (1990). Trabajo en donde se estudiaron los factores correlacionados con la transformación y la disponibilidad subsecuente de zinc en suelos calcáreos sometidos a diferentes tratamientos.

Zhang (1991). Estudio de las relaciones entre las diferentes fracciones de P en el suelo y su efecto sobre la disponibilidad de este elemento para las plantas.

Aplicaciones en Agronomía y Horticultura

Bowers *et al.* (1990). Estudio del efecto de una serie de variables ambientales, con énfasis en la precipitación pluvial, sobre la incidencia y transmisión del tizón causado por *Phytophthora capsici* en un cultivo de chile.

Karlsson *et al.* (1988). Este es un trabajo relacionado con la optimización de condiciones para la producción comercial de crisantemo (*Dendranthema grandiflora* Tzvelev). Se realizó un análisis de crecimiento; el análisis de senderos fue utilizado para estudiar la influencia del tiempo de desarrollo, y de la ganancia relativa en peso seco durante cuatro fases del desarrollo de las plantas, sobre el tiempo de desarrollo e incremento en peso seco durante fases subsecuentes.

Sun *et al.* (1990). Análisis de un modelo causal que incluye como variable de respuesta la habilidad de llenado de frutos de *Solanum melongena* L. y como variables básicas a diferentes factores ambientales (como la temperatura y la humedad del aire) que son función de variaciones estacionales.

Aplicaciones en Mejoramiento Genético de Plantas

Mebrahtu *et al.* (1991). Estudio realizado para determinar el efecto del O_3 sobre el rendimiento y sus componentes en genotipos sensibles e insensibles de frijol. El sistema causal considerado consta de siete variables básicas y el rendimiento se semilla como variable de respuesta. La lectura de este reporte es sencilla y recomendable.

Autores varios. Modelos causales de la asociación de los componentes del rendimiento con el rendimiento mismo de diferentes cultivos. La idea básica en esta clase de estudios es ofrecer alternativas para la selección indirecta o para la selección multivariada en una especie de cultivo dada. Lo que se espera conseguir como resultado es optimizar las respuestas correlacionadas con la selección realizada para alto rendimiento, evitando que dichas respuestas sean indeseables y permitir, por otro lado, construir funciones multivariadas con las cuales obtener respuestas adecuadas en un conjunto de caracteres de importancia económica.

Entre los trabajos abajo mencionados el de Adams (1967) es un escrito clásico que establece la base genética y fisiológica de la respuesta compensatoria de los componentes del rendimiento. En cuanto a las respuestas correlacionadas los trabajos de Robinson *et al*. (1951), Kempthorne (1969), Duarte y Adams (1972) y Puri *et al.* (1982) establecieron la base teórica para su manejo en los programas de mejoramiento genético de plantas. Esta área de estudio de las ciencias biológicas es en la que se encuentra mayor cantidad de aplicaciones publicadas del análisis de senderos pero asimismo es en donde se encuentra la mayor cantidad de aplicaciones que no consideran de forma adecuada la estructura causal del sistema. Un listado no exhaustivo de autores que aplican el análisis de senderos en el mejoramiento genético de plantas es:

Abraham (1990); Adams (1967); Alam *et al.* (1988); Biradar *et al.* (1991); Castinerias y Rivero (1988); Dechev

(1990); Debnath y Khan (1991); Duarte y Adams (1972); Gill *et al.* (1990); Gravois y Helms (1992); Hondelmann y Strauss (1990); Kang *et al.* (1983); Khattra *et al.* (1990); Khorgade *et al.* (1990); Madalageri y Gopinath (1990); Mahajan y Nayeem (1989); Mani y Bahl (1990); Maurya y Sinha (1987); Ombakho y Tyagi (1989); Oseni *et al.* (1992); Parh *et al.* (1990); Patil y Deshmukh (1989); Patil *et al.* (1990); Prasad y Rao (1989); Puri *et al.* (1982); Ram y Hemaprabha (1992); Reddy y Jhansilakshmi (1991); Reddy y Lal (1987); Reddy y Nerkar (1990); Robbins y Moore (1991); Sahoo *et al.* (1990); Sandhu *et al.* (1991);Sarkar (1990); Shabana *et al.* (1990); Siefker y Hancock (1986); Singh y Labana (1990); Singh, G. *et al.* (1989); Singh, R. *et al.* (1989); Solanki *el al.* (1987); Srivastava *et al.* (1988); Tagore y Singh (1990); Verma y Jha (1990).

Conclusión

De lo visto arriba es apreciable la utilidad del análisis de senderos en los estudios biológicos. Es común encontrarse con trabajos experimentales en sistemas causales bien diseñados y con buena colecta de datos, pero analizados con modelos de regresión no aplicables a la situación, o bien con técnicas multivariadas más sofisticadas, como el análisis de componentes principales, o el de factores, que tampoco puedan ser las más adecuadas a la situación.

Por otro lado, con la posible excepción de los genetistas de plantas, las técnicas de análisis estructural se encuentran al parecer poco difundidas entre los investigadores dedicados a las ciencias biológicas. Por otro lado, tal como puede apreciarse en el listado de literatura de la parte anterior, buena parte de los trabajos incluidos se refieren a situaciones que requieren cierto enfoque holístico (como en los estudios ecológicos y evolutivos citados) para lograr la comprensión y después el control de las respuestas correlacionadas en los sistemas bajo estudio.

Como cualquier herramienta el análisis de senderos es útil si se aplica correctamente. Aunque en la revisión de aplicaciones no se marcaron calificaciones de bueno o malo, la calidad de dichos reportes es bastante variable. El error menos común es que el artículo presente errores en la estimación de los coeficientes del sistema causal; el más frecuente es que el sistema causal que supuestamente define la situación tenga una solución trivial o carezca de relación con la lógica o la realidad del problema.

Por otra parte, los trabajos de buena calidad involucraron un conocimiento teórico adecuado del comportamiento biológico y de las interrelaciones de las variables que se estudiaron (lo cual no es igual a conocer únicamente la base de datos experimentales), además de una buena base de sentido común en combinación con el conocimiento de los métodos adecuados de análisis de los datos.

Literatura Citada

Adams, M.W. 1967. Basis of yield component compensation in crop plants with special reference to the field bean, *Phaseolus vulgaris.* Crop Sci. 7:505-510.

Abraham, M.J. 1990. Correlation, path and discriminant function analyses in groundnut grown on a phosphorus deficient acidic soil. Crop Improv. 17:34-37.

Abrol, D.P. 1990. Pollination activity of alfalfa-pollinating subtropical bees *Megachile nana* and *Megachile flavipes* (Hymenoptera: Megachilidae). Trop. Ecol. 31:106-115.

Alam, M.S., D.K. Parh and M.Z.I. Talukder. 1988. Correlation and path-coefficient analysis in sunflower (*Helianthus annuus* L.). Bangladesh J. Agric. 13:147-150.

Asher, H.B. 1983. Causal Modeling. Sage Publications, Inc. Newbury Park, California. U.S.A.

Biradar, B.D., J.V. Goud and S.S. Patil. 1991. A study on character association and path coefficient in cowpea. J. Maharashtra Agric. Univ. 16:27-29.

Bowers, J.H., R.M. Sonoda and D.J. Mitchell. 1990. Path coefficient analysis of the effect of rainfall variables on the epidemiology of phytophthora blight of pepper caused by *Phytophthora capsici.* Phytopathology 80:1439-1446.

Castinerias, L. y N. Rivero. 1988. Análisis de correlación y de coeficientes de sendero para el rendimiento y sus componentes en el frijol común. Cienc. Agric. 0(34/35):77-81.

Chanter, D.O. 1981. The use and misuse of linear regression methods in crop modelling. In: D.A. Rose and D.A. Charles-Edwards (Eds.). Mathematics and Plant Physiology. Academic Press, Inc. (London) Ltd. London, U.K. p. 253-267.

Crespi, B.J. 1990. Measuring the effect of natural selection on phenotypic interaction systems. Am. Nat. 135:32-47.

Crespi, B.J. and F.L. Bookstein. 1988. A path-analytic model for the measurement of selection on morphology. Evolution 43:18-28.

Debnath, S.C. and M.F. Khan. 1991. Genotypic variation, covariation and path coefficient analysis in maize. Pak. J. Sci. Ind. Res. 34:391-394.

Dechev, D. 1990. Path analysis of yield components in durum wheat. Genet. Sel. 23:213-216.

Dofing, S.M. and C.W. Knight. 1992. Alternative model for path analysis of small-grain yield. Crop Sci. 32:487-489.

Draper, N.R. and H. Smith. 1981. Applied regression analysis. John Wiley & Sons, Inc. New York, U.S.A.

Duarte, R.A. and M.W. Adams. 1972. A path coefficient analysis of some yield component interrelations in field bean (*Phaseolus vulgaris* L.). Crop Sci. 12:579-582.

Farris, M.A. and M.J. Lechowicz. 1990. Functional interactions among traits that determine reproductive success in a native annual plant. Ecology 71:548-557.

Ferrari, T.J. 1963. Causal soil-plant relationships and path coefficients. Plant Soil 13:81-96.

Gill, K., T.H. Singh and G.S. Chahal. 1990. Path analysis of genotype-environment interactions in upland cotton. Crop Improv. 17:27-30.

Gravois, K.A. and R.S. Helms. 1992. Path analysis of rice yield and yield components as affected by seeding rate. Agron. J. 84:1-4.

Hoerl, A.E. and R.W. Kennard. 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12:55-67.

Hondelmann, W. and D.D. Strauss. 1990. Path-coefficient analysis of seed yield components in *Euphorbia lathyris.* Plant Breeding 105:112-116.

Jordan, N. 1992. Path analysis of local adaptation in two ecotypes of the annual plant *Diodia teres* Walt. (Rubiaceae). Am. Nat. 140:149-165.

Kang, M.S. 1992. Letter to the editor. Agron. J. 84:917-918.

Kang, M.S., J.D. Miller and P.Y.P. Tai. 1983. Genetic and phenotypic path analyses and heritability in sugarcane. Crop Sci. 23:643-647.

Karlsson, M.G., M.P. Pritts and R.D. Heins. 1988. Path analysis of chrysanthemum growth and development. HortScience 233:372-375.

Kempthorne, O. 1969. An Introduction to Genetic Statistics. The Iowa State University Press. Ames, Iowa. U.S.A.

Kendall, M. and A. Stuart. 1976. The Advanced Theory of Statistics. Vol. 3. Charles-Griffin & Co. Ltd. London, U.K.

Khattra, A.S., K.S. Nandpuri and J.C. Thakur. 1990. Correlation and path coefficient analysis in tomato. J. Res. Punjab Agric. Univ. 27:412-418.

Khorgade, P.W., A.H. Nafade, M.N. Narkhede and S.K. Raut. 1990. Some selection criteria in green gram. J. Maharashtra Agric. Univ. 15:179-182.

Leduc, A., P. Drapeau, Y. Bergeron and P. Legendre. 1992. Study of spatial components of forest cover using partial Mantel test and path analysis. Veg. Sci. 3:69-78.

Li, C.C. 1956. The concept of path coefficient and its impact on population genetics. Biometrics 12:191-210.

Li, C.C. 1968. Fisher, Wright and path coefficients. Biometrics 24:471-483.

Li, C.C. 1975. Path Analysis -A Primer. The Boxwood Press.Pacific Grove, California. U.S.A.

MacDonald, K.I. 1977. Path analysis. In: C.A. O'Muircheartaigh and C. Payne (Eds.). The Analysis of Survey Data (Vol. 2). John Wiley & Sons, New York. U.S.A.

Madalageri, B.B. and G. Gopinath. 1990. Correlation and path coefficient analysis for dry matter content in half-sib families of Bellary red onion (*Allium cepa* L.). Mysore J. Agric. Sci. 22:479-482.

Mahajan, A.R. and K.A. Nayeem. 1989. Character association and path analysis for seedling vigor index in wheat species and triticale under different temperatures. Seed Res. 17:32-35.

Mani, M. and P.N. Bahl. 1990. Components of productivity in contrasting types of chickpea (*Cicer arietinum* L.). J. Genet. Breed. 44:163-168.

Maurya, K.R. and R.P. Sinha. 1987. Path coefficient analysis in fenugreek (*Trigonella foenum graceum* L.). Prog. Hortic. 19:276-279.

May, L., D.A. Van Sanford, C.T. MacKown and P.L. Cornelius. 1991. Genetic variation for nitrogen use in soft red x hard red winter wheat populations. Crop Sci. 31:626-630.

Mebrahtu, T., W. Mersie and M. Rangappa. 1991. Path coefficient analysis of ozone effects on seed yield and seed yield components of bean (*Phaseolus vulgaris* L.). J. Hortic. Sci. 66:59-66.

Mitchell, R.J. 1992. Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. Funct. Ecol. 6:123-129.

Mitchell-Olds, T. 1987. Analysis of local variation in plant size. Ecology 68:82-87.

Mitchell-Olds, T. and J. Bergelson. 1990. Statistical genetics of an annual plant, *Impatiens capensis*. II. Natural selection. Genetics 124:417-421.

Ombakho, G.A. and A.P. Tyagi. 1989. Correlation and path-coefficient analysis for yield and its components in cowpea (*Vigna unguiculata* (L.) Walp). East Afr. Agric. For. J. 53:23-28.

Oseni, T.O., D.D. Lenge and U.R. Pal. 1992. Correlation and path-coefficient analysis of yield attributes in diverse lines of cowpea (*Vigna unguiculata*). Indian J. Agric. Sci. 52:365-368.

Pantone, D.J., W.A. Williams and A.R. Maggenti. 1989. An alternative approach for evaluating the efficacy of potential biocontrol agents of weeds: 2. Path analysis. Weed Sci. 37:778-783.

Patil, H.S. and R.B. Deshmukh. 1989. Correlation and path analysis in black gram. J. Maharashtra Agric. Univ. 14:310-312.

Patil, H.S., B.N. Narkhede and D.R. Bapat. 1990. Association and path analysis in sorghum. J. Maharashtra Agric. Univ. 15:18-19.

Pedhazur, E.J. 1982. Multiple Regression in Behavioral Research. Holt, Rinehart and Winston, New York, U.S.A.

Pindyck, R.S. and D.L. Rubinfeld. 1981. Econometric Models and Economic Forecasts. McGraw-Hill Book Co. Singapore. Singapore.

Prasad, B., A.K. Metha and M.K. Sinha. 1990. Zinc fractions and availability of applied zinc in calcareous soil treated with organic materials. J. Indian Soc. Soil Sci. 38:248-253.

Prasad, M.B.N.V. and G.S.P. Rao. 1989. Genetic variability, correlations and path-coefficient analysis for some morphological and biochemical constituents of acid lime fruit. Sci. Hortic. 41:43-54.

Pritts, M.P. and J.F. Hancock. 1985. Lifetime biomass partitioning and yield component relationships in the highbush blueberry, *Vaccinium corymbosum* L. (Ericaceae). Amer. J. Bot. 72:446-452.

Puri, Y.P., C.O. Qualset and W.A. Williams. 1982. Evaluation of yield components as selection criteria in barley breeding. Crop Sci. 22:927-931.

Ram, B. and G. Hemapraba. 1992. Character interrelationships in cultivar x species progenies in sugarcane. Indian J. Genet. Plant Breed. 51:89-95.

Reddy, C.D.R. and K. Jhansilakshhmi. 1991. Variability and path analysis of component characters in foxtail millet. J. Maharashtra Agric. Univ. 16:44-47.

Reddy, C.D.R. and Y.S. Nerkar. 1990. Studies on correlations and path analysis in parental lines sensitive to iron chlorosis and F_1s (tolerant x sensitive) in rice (*Oryza sativa* L.) under upland irrigated conditions. Crop Res. (HISAR) 3:61-65.

Reddy, M.L.N. and G. Lal. 1987. Genetic variability and path coefficient analysis in tomato (*Lycopersicon esculentum* Mill.) under summer season. Prog. Hortic. 19:284-288.

Reyes, M.H. 1992. Análisis de Senderos con el Paquete MATLAB. Laboratorio de Genética Biométrica. Universidad Autónoma Agraria Antonio Narro. Saltillo, México.

Reyes, M.H. y A. Benavides. 1993. El Análisis de Senderos en la Investigación Agronómica. Universidad Autónoma Agraria Antonio Narro. Folleto de Divulgación Vol. III (4).

Robbins, J. and P.P. Moore. 1991. Fruit morphology and fruit strength in a seedling population of red raspberry. HortScience 26:294-295.

Robinson, H.F., F.E. Comstock and P. Harvey. 1951. Genotypic and phenotypic correlations in corn and their implications in selection. Agron. J. 43:282-287.

Sahoo, S.C., S.N. Mishra and R.S. Mishra. 1990. Association of traits in segregating populations of chilli (*Capsicum annuum*). Indian J. Agric. Sci. 60:233-236.

Sandhu, T.S., R.K. Gumber and R.S. Bhatia. 1991. Path analysis in chickpea. J. Res. Punjab Agric. Univ. 28:1-4.

Sarkar, R.K. 1990. Correlation and path analysis of certain morphophysiological characters in soybean. Indian J. Plant Physiol. 33:83-84.

Schemske, D.W. and C.C. Horvitz. 1988. Plant-animal interactions and fruit production in a neotropical herb: a path analysis. Ecology 69:1128-1137.

Shabana, R., S.A. Shhries, A.F. Ibrahim and G. Geisler. 1990. Correlation and path coefficient analysis for some new released spring rapeseed cultivars grown under different competitive systems. J. Agron. Crop Sci. 165:138-143.

Shasha'a, N.S., W.P. Nie and W.F. Campbell. 1973. Path-coefficient analysis of correlation between honey bee activity and seed yield in *Allium cepa* L. J. Amer Soc. Hort. Sci. 98:341-347.

Siefker, J.H. and J.F. Hancock. 1986. Yield component interactions in cultivars of the highbush blueberry. J. Amer. Soc. Hort. Sci. 111:606-608.

Singh, G., H.S. Brar, M.M. Verma and J.S. Shandu. 1989. Component analysis of seed yield in chickpea. Crop Improv. 16:145-149.

Singh, R., B.S. Sukhija and J.S. Hundal. 1989. Correlation and path coefficient studies in carrot. J. Res. Punjab Agric. Univ. 26:581-584.

Singh, S.B. and K.S. Labana.1990. Correlation and path analysis in sunflower. Crop Improv. 17:49-53.

Solanki, S.S., P.K. Saxena and I.C. Pandey. 1987. Path analysis of yield components in pea (*Pisum sativum* L.). Prog. Hortic. 19:268-270.

Srivastava, H.M., B.L. Srivastava, M.P. Agarwal and D.V.K. Saxena. 1988. Components of recoverable sugar and path analysis in diploid varieties of sugar beet. Indian J. Genet. Plant Breed. 49:403-406.

Stewart, G.W. 1987. Collinearity and least squares regression. Statistical Science 2:68-100.

Sun, W.Q. 1989. Path analysis of relationship between the endogenous polyamines and lateral root initiation in apple seedlings. Acta Phytophysiol. Sin. 15:388-392.

Sun, W.Q., D. Wang, Z. Wu and J. Zhi. 1990. Seasonal change of fruit setting in eggplants (*Solanum melongena* L.) caused by different climatic conditions. Sci. Hortic. 44:55-60.

Tagore, K.R. and I.S. Singh. 1990. Character association and path analysis under two levels of management in chickpea. Crop Improv. 17:41-44.

Talukdar, F.A., M. Shah Jahan and S. Ahmed. 1991. Effect of nitrogen and zinc on rice green leafhoppers, *Nephotettix* spp. and prey-predator relationship between leafhoppers and their predators. Pak. J. Sci. Ind. Res. 34:453-455.

Turner, M.E. and C.D. Stevens. 1959. The regression analysis of causal paths. Biometrics 15:236-258.

Verma, V.S. and V.B. Jha. 1990. Correlation and path analysis for quality traits in potato (*Solanum tuberosum* L.). J. Indian Potato Assoc. 17:169-172.

Wesser, S.D. and W.S. Armbruster. 1991. Species distribution controls across a forest-steppe transition: a causal model and experimental test. Ecol. Monographs. 61:323-342.

Wright, S. 1921. Correlation and causation. J. Agric. Res. 20:557-587.

Wright, S. 1934. The method of path coefficients. Annals of Mathematical Statistics 5:161-215.

Wright, S. 1960. Path coefficients and path regressions: alternative or complementary concepts? Biometrics 16:189-202.

Zhang, W. 1991. Path analysis and relative availability of inorganic and organic phosphorus fractions in soils. Acta Pedol. Sin. 28:4177-425.