

**AGRUPAMIENTO DE PRODUCTOS AGRÍCOLAS MEDIANTE EL ANÁLISIS DE
PLAGAS Y ENFERMEDADES COMO METODOLOGÍA PARA FACILITAR LA
GESTIÓN DE INSUMOS.**

**AGRICULTURAL PRODUCTS CLUSTERING THROUGH THE PESTS AND
DISEASES ANALYSIS AS METHODOLOGY TO IMPROVE THE INPUTS
MANAGEMENT.**

Leonardo Hernán Talero Sarmiento (UNIVERSIDAD INDUSTRIAL DE SANTANDER)

leonardo.talero@corro.uis.edu.co

Laura Yeraldín Escobar Rodríguez (UNIVERSIDAD INDUSTRIAL DE SANTANDER)

laura.escobar@correo.uis.edu.co

Edwin Alberto Garavito Hernández (UNIVERSIDAD INDUSTRIAL DE SANTANDER)

garavito@uis.edu.co

Resumen

El objetivo del presente trabajo es plantear una metodología para la identificación de insumos afines requeridos en la producción agrícola, mediante el agrupamiento de diferentes productos cultivados en una región a partir analizando similitud de afectación por plagas y enfermedades. La metodología planteada es aplicada en un escenario compuesto por cultivos presentes en Santander, Colombia, en donde a partir de la estimación del coeficiente de similitudes, se realiza la agrupación de los productos. Esta agrupación permite afirmar relaciones naturales entre algunos cultivos, así como plantear nuevas familias de productos no tan evidentes, lo cual se traduce en la identificación de requerimientos de insumos para modelos mixtos de producción.

Palabras Claves: Producción agrícola, coeficientes de similitud, coeficiente de Jaccard, Agrupamiento.

Abstract

The objective of the present work proposes a methodology for inputs identifications in agricultural production resource planning, by clustering different products grown in a specific region, based on the of the similarity coefficient estimation (for pests and diseases). The methodology was applied in Santander Colombia, with its specific food products, as from the similarity coefficient the clustering of the products is performed. This clustering allows affirm

natural relations between some crops, as well as to propose new families of products (not so obvious) which turn in the identification of input requirements for mixed production models.

Key-words: Agricultural production, similarity coefficient, Jaccard coefficient, Clustering.

1. Introducción

Existen numerosos factores que pueden ocasionar disminución en la eficiencia de la producción agrícola o la pérdida parcial o total de los cultivos, entre esos factores, se encuentra la falta de planeación para contrarrestar plagas y enfermedades propias de los productos a cultivar. Para el desarrollo de un potencial productivo, es necesario brindar desde la academia herramientas o elementos para que los pequeños y grandes productores puedan realizar una adecuada planeación de la producción y así evaluar previamente los recursos a utilizar en todo el ciclo productivo, evitando de esta manera incurrir en sobre costos, desperdicios o pérdidas.

El objetivo de la presente investigación es presentar una metodología para identificar características agrícolas similares con el fin de mejorar la planeación de los recursos o insumos a utilizar durante la producción mediante la creación de grupo o familias de productos, teniendo en cuenta las plagas y enfermedades en común. La metodología es evaluada con una serie de cultivos propios de Santander, Colombia, teniendo en cuenta que, existe una alta biodiversidad y potencial productivo.

El presente documento se compone de un primer capítulo denominado *Planteamiento del problema*, en el que se realiza un marco de antecedentes de los diferentes coeficientes de similitud existentes y sus aplicaciones, continuando con un acercamiento a la necesidad de generar propuestas para el desarrollo agrícola de Santander; un segundo capítulo denominado *Metodología* se presenta a continuación, en donde se explican las etapas de desarrollo de la propuesta de planeación realizada a partir de la clusterización de cultivos de la región, en el capítulo de *Resultados* se consignan las relaciones estimadas a partir del grado de similitud, distancia entre productos y agrupamiento, finalmente: *Discusiones, Conclusiones y Bibliografía*.

2. Marco de antecedentes

Las relaciones fenéticas o de similitud, se basan en el parecido entre dos o más organismos, teniendo en cuenta las propiedades observadas en ellos, sin considerar el proceso genealógico por el cual aparecieron estas propiedades (CRISCI; LÓPEZ, 1983). Buscando cuantificar dichas relaciones, se han propuesto coeficientes o índices de similitud o similaridad, que a partir de

datos cualitativos o cuantitativos permiten comparar propiedades o estados de dos o más comunidades. Crisci y López, determinan tres grandes grupos de coeficientes: de distancia, de correlación y de asociación. Los coeficientes de distancia que busca determinar a partir de la distancia existente entre dos entidades, la similitud o disimilitud que existe entre ellos, a mayor distancia, menor grado de similitud, ejemplo de éstos es el coeficiente de distancia promedio o Mean Character Distance (MCD) por sus siglas en inglés, propuesto por Cain y Harrison (CAIN; HARRISON, 1958). Los coeficientes de correlación permiten cuantificar la similitud de entidades, a partir de la medición de la separación angular formada por las líneas que parten de un origen de coordenadas y pasan por las entidades a comparar, Michener y Sokal trabajaron con el coeficiente de Pearson o el también llamado coeficiente de correlación de momento-producto (SOKAL, 2015). Por otra parte, los coeficientes de asociación miden las coincidencias y diferencias en los diferentes estados de dos o más comunidades o entidades, para dicha medición, se requieren datos del tipo doble – estado (presencia/ ausencia) o binarios, es decir, datos que únicamente tienen dos estados que a su vez son excluyentes entre sí. Existen numerosos coeficientes de asociación, entre ellos los propuestos Jaccard, Sorensen y Dice, Roger y Tanimoto, Sokal y Sneath, Hamann, entre otros, tal como los menciona Crisci y López (1983).

Estos coeficientes han reflejado su utilidad en diversos campos de la ciencia. En biología, se han utilizado criterios de similaridad para analizar comunidades, así como la organización de las unidades ecológicas, bien sea dentro de una misma unidad o entre unidades (SAIZ, 1980). Otro campo en el que se han desarrollado aplicaciones para estos índices es la ciencia forense en donde se realizan estas relaciones con el fin de determinar patrones de conducta de criminales (BENNELL; GAUTHIER; GAUTHIER; MELNYK; MUSOLINO, 2010), en la industria química se han realizado análisis de las distribuciones de los valores de similitud intermoleculares a partir de los coeficientes de Tanimoto y de coseno (HOLLIDAY; SALIM; WILLETT, 2005). En la industria de manufactura, también se han realizado estudios con el fin de abordar el problema de formación de células de manufactura a partir de coeficientes de similaridad, tales son los casos de Dimopoulos y Mort (2000) quienes realizan un estudio de la evolución de los índices de similitud empleados en el contexto; Yin y Yasuda, a su vez, plantean una investigación comparativa de los métodos de coeficientes de similitud aplicados al problema de formación de células (YIN; YASUDA, 2006); entre otros.

El índice de Jaccard es uno de los coeficientes más usados en la literatura para evaluar la similitud que existe entre dos o más comunidades, gracias a las características matemáticas y

estadísticas que posee. Existen numerosas aplicaciones del índice de Jaccard, en biología se ha utilizado para comparar especies (WILCHES, 1995); en biología molecular se ha revisado la utilidad de los coeficientes de similitud para realizar análisis de agrupamiento (DEN; DEPAUW; MOM ILOVI; A IN, 2010); en ingeniería, se han registrado aplicaciones para buscar eficiencia en producción a partir del agrupamiento de máquinas (MCAULEY, 1972), entre otros.

Partiendo de la transversalidad de estos coeficientes en distintas áreas, incluyendo la producción, se plantea la posibilidad de usar los índices para determinar el grado de similitud de distintos productos en la producción agrícola, evitando así pérdidas de productos. Contextualizando dicha problemática, en los últimos años, la sociedad de Agricultores de Colombia ha reportado una caída en el valor de la producción agropecuaria, derivada principalmente a la contracción de las actividades agrícolas de ciclo corto, así como la variación de los precios, las condiciones climáticas, sobrecostos de insumos, entre otros factores, que han ocasionado pérdidas en las siembras, disminución en rendimientos por hectárea y baja calidad en los productos (MEJÍA, 2016). De acuerdo a lo anterior y entendiendo que Colombia y en particular Santander, por su ubicación estratégica, así como por las bondades propias del territorio, puede desarrollar en mayor proporción el sector agrícola, se hace necesario revisar alternativas que permitan a los pequeños y grandes productores producir de manera eficiente sus cultivos, planeando el volumen de producción, logística de abastecimiento de insumos; minimizando de esta manera el riesgo de pérdida de productos por condiciones propias del ciclo productivo, como lo son las plagas y enfermedades de los productos. Partiendo de estos hechos, se propone realizar una organización de un mix de productos que puedan cultivarse en el territorio santandereano y que a partir de la similitud que tienen estos en cuanto a recursos, se puedan estimar las necesidades de la producción para que una plaga o una enfermedad no tome por sorpresa a los productores y termine disminuyendo la eficiencia del cultivo realizado.

3. Metodología

El presente trabajo de investigación tiene como objetivo desarrollar una metodología para brindar herramientas para la realización de la planeación de los recursos a utilizar en la producción agrícola, tomando como caso productos cultivados en Santander, Colombia; enfocado a gestionar el acopio de insumos (plaguicidas y demás recursos existentes para el control de plagas y enfermedades propias de los productos del departamento) a partir de la identificación de familias. La metodología planteada se divide en cinco etapas:

- Identificación de plagas y enfermedades propias de productos en Santander.
- Selección del coeficiente de similitud
- Construcción de la matriz de coeficientes de similitud.
- Construcción de la matriz de distancia
- Clusterización de cultivos.

3.1. Identificación de plagas y enfermedades propias de productos en Santander

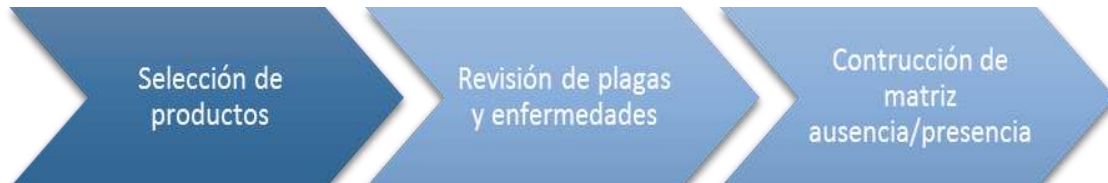
Para esta primera etapa, se hace una revisión de los cultivos propios del departamento de Santander y a partir de ésta, la selección 18 productos para los cuales se identificaron las plagas y enfermedades. A continuación, se presenta el listado de los productos objeto de estudio.

- Acelga (CANCINO, 2012)
- Arracacha amarilla (DANE, 2015b)
- Arveja verde en vaina (DANE, 2015a)
- Banano criollo (LEÓN; GUALDRÓN, 2010)
- Cebolla Cabezona blanca (DANE, 2016b)
- Cebolla junca (DANE, 2015c)
- Fríjol verde en vaina (ARIAS; RENGIFO; JARAMILLO, 2007)
- Habichuela (DANE, 2016a)
- Limón Tahití (SIPSA, 2015)
- Lulo (DANE y MINAGRICULTURA, 2014)
- Mora de castilla (MINAGRICULTURA y DANE, 2013)
- Papa (DANE, 2013a)
- Papa criolla limpia (PIÑEROS, 2009)
- Papaya maradol (DANE, 2016c)
- Pimentón (DANE, 2013b)
- Tomate riñón (DANE, 2014)
- Yuca criolla (BANCO INTERAMERICANO DE DESARROLLO, 2005)
- Zanahoria (NÚCLEO AMBIENTAL S.A.S, 2015)

Para la construcción de este listado, se discriminaron aspectos como el tipo de producto (hortalizas, tubérculos, frutas, entre otros), tipo de cultivo (transitorio/permanente), ciclo de producción, costos de producción y precios de comercialización. Seguido a esto y con base en la información encontrada en las fuentes anteriores, se elaboró una matriz con el fin de

representar el factor o estado (plagas y enfermedades) en común que tenía cada par de cultivos, para esto se valoró con 1 la presencia del factor a evaluar en dos cultivos y con 0 la ausencia del mismo. Una vez realizada la revisión de enfermedades y plagas presentes en el total de cultivos, se identifican 104 factores a clasificar y evaluar para cada par de cultivos.

Gráfica 1. Proceso de identificación de plagas y enfermedades en cultivos de Santander



3.2. Selección del coeficiente de similitud

Partiendo del planteamiento del problema, se determina que se revisará únicamente la presencia de los factores o estados (enfermedades y plagas) en común en dos comunidades y se discriminará como criterio de similitud los factores ausentes en las mismas comunidades. De acuerdo a lo anterior, se plantea hacer uso del coeficiente de Jaccard (MOSTACEDO; FREDERICKSEN, 2000) expresado de la siguiente manera:

$$IJ = \frac{C}{A + B - C} * 100$$

Donde:

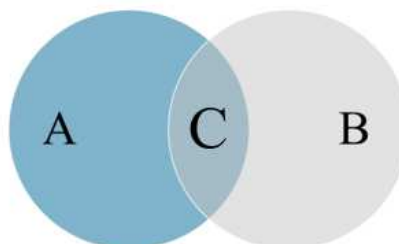
IJ: Índice o coeficiente de Jaccard.

A: Número de especies de enfermedades y plagas presentes en la comunidad *A*.

B: Número de especies de enfermedades y plagas presentes en la comunidad *B*.

C: Número de especies de enfermedades y plagas comunes en *A* y *B*.

Gráfica 2. Criterios del Coeficiente de Jaccard



Este índice permitirá conocer la porción de enfermedades y plagas en común que tienen cada par de productos evaluados, respecto a la cantidad global de enfermedades y plagas que tienen en total dichos productos. Adicional a lo anterior y de acuerdo con Gower (1966) el coeficiente de similitud propuesto por Sokal y Michener, así como el propuesto por Jaccard, dan como resultados matrices de similaridad definidas semi positivas, condición que permite calcular la distancia a partir de dichos coeficientes.

3.3. Construcción de la matriz de coeficientes de similitud

Con el objetivo de conocer la similitud existente bajo el criterio de plagas y enfermedades presentes en cada par de cultivos, se realizan los cálculos pertinentes, siguiendo la expresión definida en el numeral anterior. En esta etapa se comparan todas las plagas y enfermedades presentes en los 18 productos seleccionados y se determina el porcentaje de similitud entre cada par de productos, dando como resultado una matriz de 18x18.

3.4. Construcción de la matriz de distancia

Existen numerosas expresiones que relacionan los coeficientes de similitud con la distancia; sin embargo, para la presente aplicación se evaluará la distancia euclidiana entre dos entidades de la siguiente forma:

$$D_{ij}^2 = 2(1 - S_{ij})$$

Donde:

D_{ij}: Distancia euclidiana entre las entidades *ij*.

S_{ij}: Coeficiente de similitud entre las entidades *ij*.

De acuerdo con Wilches (1995), Gower hizo énfasis en utilizar la expresión anteriormente mencionada y calcular su resultado a partir de los coeficientes de similitud de Sokal y Michener o de Jaccard, debido a que estos arrojan matrices definidas como semi positivas, hecho que facilita convertir las similitudes a distancias.

3.5. Clusterización de cultivos

Una vez obtenida la matriz de distancia, se realiza un agrupamiento mediante una técnica jerárquica en el software estadístico R. Las técnicas de agrupación permiten clasificar las observaciones similares en grupos y para el caso de los métodos jerárquicos, éstos presentan una mayor exactitud que los métodos de permutación como el k-means (KARYPIS; KUMAR;

STEINBACH, 2000) (LI; CHU; WANG; YANG, 2002), sin embargo, el método de enlace y la medida de distancia influye considerablemente en los resultados finales.

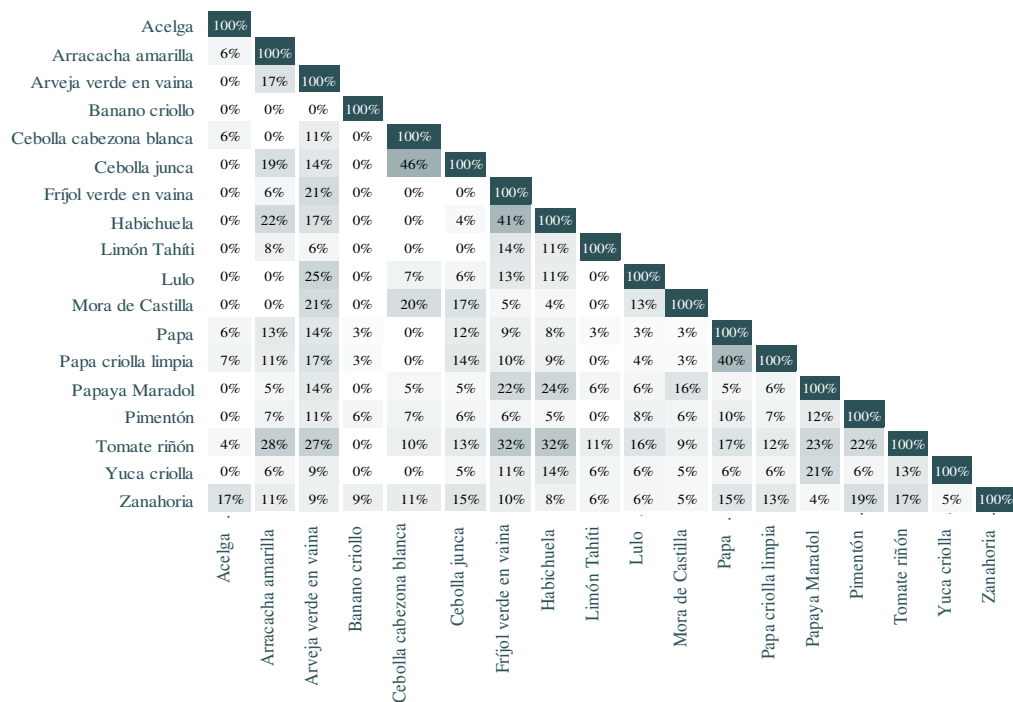
Teniendo en cuenta lo anterior y los resultados obtenidos por Galili (2017), se determina trabajar con un método de enlace Ward debido a su similitud con los demás tipos de enlace, además, buscando minimizar la suma de cuadrados dentro del conglomerado. Teniendo en cuenta el tipo de enlace, se escoge la función de distancia de Pearson cuadrada, ésta es la suma de las distancias elevadas al cuadrado dividida entre las varianzas, constituyendo así una medida estandarizada de la distancia euclídea definida en el numeral anterior.

Para determinar el número máximo de grupos, se analizó la distancia entre conjunto a partir de la creación iterativa de clústeres, identificando el valor de la Suma Cuadrada de la Distancia (SCD) entre los grupos, se crearon un total de 10 escenarios, y se escogió aquel donde el cambio de SCD era despreciable. (KASSAMBARA; MUNDT, 2017).

4. Resultados

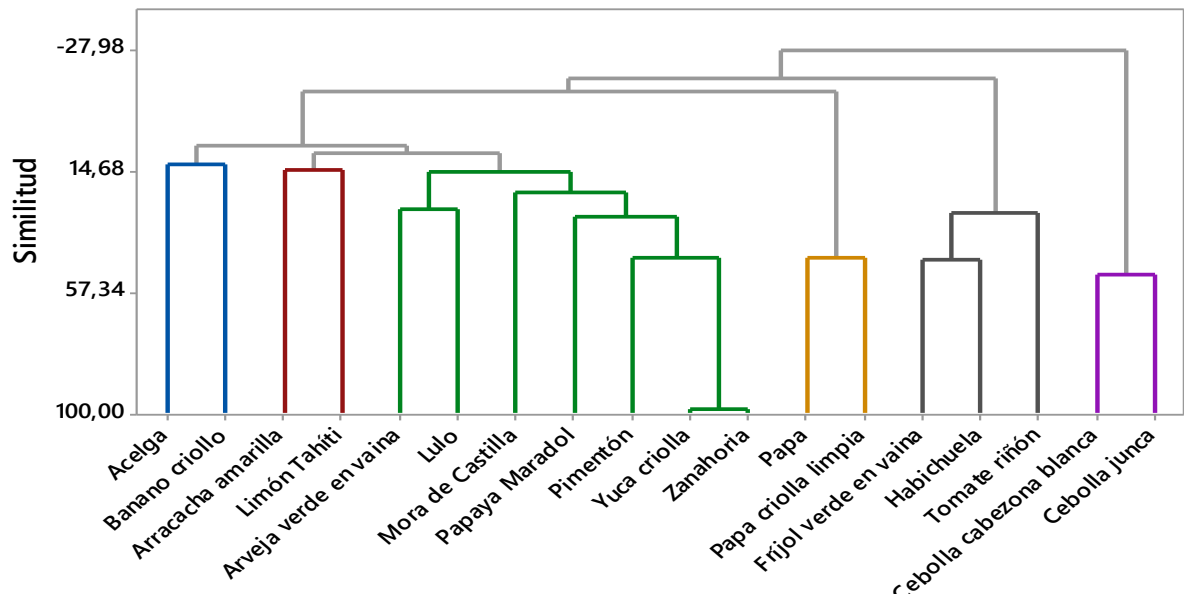
A partir de la selección de los 18 productos ya mencionados, se revisa e identifica por medio de una matriz la presencia/ausencia (1/0) de 104 enfermedades y plagas en cada uno de esos cultivos. Se realizó la evaluación de los coeficientes de similitud de 18 productos, lo que arrojó una matriz de 18x18, en la Gráfica 3 se presentan índices de similitud entre pares de cultivos tomando como principal criterio las plagas y enfermedades en común.

Gráfica 3. Matriz de Coeficientes de Similitud



distancias atípicas entre ellos, es decir, poseen un valor de lejanía mayor al esperado, lo cual discurre en sus características o factores en común, siendo poco parecidos al resto y posiblemente más resistentes a las plagas y enfermedades.

Gráfica 6. Dendograma de los productos cultivados en Santander según afinidad en plagas y enfermedades



5. Discusión

El presente trabajo plantea una metodología para la planeación de los recursos de la producción agrícola a partir de la clusterización de cultivos, utilizando como principal criterio de agrupación la similitud en cuanto a plagas y enfermedades entre pares de productos. Si bien la literatura provee amplia información de las aplicaciones de coeficientes de similitud en áreas como la química, medicina, biología, ingeniería, entre otras; no se evidencian numerosos estudios de la aplicación de estos índices para la planeación de la producción agrícola.

En el presente estudio no se tuvo en cuenta factores importantes de la producción agrícola tales como el tipo de cultivo, el suelo del territorio de cultivo, costos de producción, precios de comercialización entre otros, debido a esto, se sugiere evaluar la metodología utilizando otros criterios de similitud para analizar su validez en otros escenarios. Adicional a lo anterior, se sugiere extender la metodología hasta el cálculo de los recursos necesarios para cada clúster, con el fin de realizar un contraste entre los costos totales de insumos actualmente y los costos proyectados en dicho estudio.

Es necesario además, evaluar el resultado entregado en el presente estudio, con el resultado producto del uso de otros coeficientes de similitud, por ejemplo el coeficiente planteado Sokal

y Michener, y de esta manera conocer si el coeficiente utilizado para la presente metodología es el mejor para estos casos o no.

6. Conclusiones

El potencial agrícola productivo en una región puede estar comprometido y diezmado como consecuencia de la falta de planeación de la producción, por tanto, es necesario generar herramientas y metodologías que faciliten identificar posibles riesgos derivados de la producción intensiva, si bien esta suele asociarse a los monocultivos; mediante el estudio desarrollado en la presente investigación se determina que diversos productos pueden ser nichos de plagas y enfermedades comunes fungiendo como un monocultivo.

Identificar dichas similitudes es fundamental para una correcta planeación de la producción, puesto que es posible mejorar la logística de aprovisionamiento de insumos para combatir posibles plagas y enfermedades futuras.

7. Bibliografía

ARIAS, Jesús; RENGIFO, Teresita; JARAMILLO, Maribel. **BUENAS PRÁCTICAS AGRÍCOLAS (BPA)**. 2007.

BANCO INTERAMERICANO DE DESARROLLO. **El cultivo de la yuca. 2005**, (Iic), 1-12. Recuperado de: <http://www.iadb.org/es/acerca-del-bid/acerca-del-banco-interamericano-de-desarrollo,5995.html%5Cnhttp://www.iadb.org/es/proyectos/proyectos,1229.html>

BENNELL, Craig; GAUTHIER, Donna; GAUTHIER, Donald; MELNYK, Tamara; MUSOLINO, Evanya. **The impact of data degradation and sample size on the performance of two similarity coefficients used in behavioural linkage analysis**. 2010. Forensic Science International, 199(1-3), 85-92. doi:10.1016/j.forsciint.2010.03.017

CAIN, A. J.; HARRISON, G. A. . **An Analysis of the Taxonomist'S Judgment of Affinity**. 1958. Proceedings of the Zoological Society of London, 131(1), 85-98. doi:10.1111/j.1096-3642.1958.tb00634.x

CANCINO, Janeth Morales. **El cultivo de la acelga (Beta vulgaris var. cicla)**. 2012.

CRISCI, Jorge Víctor; LÓPEZ, María Fernanda. **Introducción a la Teoría y Práctica de la Taxonomía Numérica**. 1983, 128.

DANE. **El cultivo de la papa, Solanum tuberosum Alimento de gran valor nutritivo, clave en la seguridad alimentaria mundial**. 2013. Boletín Mensual Insumos y factores asociados a la producción agropecuaria, 15, 92.

DANE. **El cultivo de pimentón bajo invernadero**. 2013. Boletín mensual, insumos y factores asociados a la producción agropecuaria, 74. Recuperado de: http://scholar.google.es/scholar?q=pistacho+cultivo+españa&btnG=&hl=ca&as_sdt=0,5#5

DANE. **El cultivo del tomate de mesa bajo invernadero, tecnología que ofrece mayor producción, calidad e inocuidad del producto**. 2014, 72. Recuperado de: http://www.dane.gov.co/files/investigaciones/agropecuario/sipsa/insumos_factores_de_produccion_dic_2014.pdf

DANE. **Boletín mensual. Insumos y factores asociados a la producción agropecuaria**. 2015, 78. Recuperado de: https://www.dane.gov.co/files/investigaciones/agropecuario/sipsa/Bol_Insumos31_mar_2015.pdf

DANE. **El cultivo de la arracacha bondades nutricionales y de alta potencialidad agroindustrial**. 2015.

- DANE. **La cebolla de rama o cebolla junca, una hortaliza de gran importancia en la alimentación humana.** 2015.
- DANE. **Cultivo de la habichuela y el fenómeno del niño.** 2016.
- DANE. **El cultivo de la cebolla cabezona frente a condiciones de alta humedad.** 2016.
- DANE. **El cultivo de papaya y sus principales enfermedades en época de lluvias.** 2016.
- DANE; MinAgricultura. **El cultivo del lulo (Solanum quitoense), una fruta agradable y de gran valor nutritivo.** 2014. Boletín mensual (SIPSA), 23, 64.
- DEN, Srbislav; DEPAUW, Ron; MOM ILOVI, Vojislava; A IN, Vladimir. **Comparison of Similarity Coefficients Used for Cluster Analysis Based on SSR Markers in Sister Line Wheat Cultivars.** 2010. Genetics and molecular research : GMR, 9(4), 2248-2253. DOI: 10.4238/vol9-4gmr966
- DIMOPOULUS, Christos; MORT, Neil. **Evolving similarity coefficients for the solution of cellular manufacturing problems.** 2000. Proceedings of the Congress on Evolutionary Computation CEC 2000, 617-624. doi:10.1109/CEC.2000.870355
- GALILI, Tal. **Hierarchical cluster analysis on famous data sets - enhanced with the dendextend package.** 2017.
- GOWER, J. C. . **Some distance properties of latent root and vector methods used in multivariate analysis.** 1966. Biometrika, 53(3-4), 325-338. doi:10.1093/biomet/53.3-4.325
- HOLLIDAY, JD; SALIM, Naomie; WILLETT, Peter. **On the magnitudes of coefficient values in the calculation of chemical similarity and dissimilarity.** 2005. American Chemical Society Recuperado de: <http://pubs.acs.org/doi/abs/10.1021/bk-2005-0894.ch006>
- KARYPIS, MSG; KUMAR, V.; STEINBACH, M. . **A comparison of document clustering techniques.** 2000. TextMining Workshop at.
- KASSAMBARA, Alboukadel; MUNDT, Fabian. **Package «factoextra»: Extract and Visualize the Results of Multivariate Data Analyses.** 2017.
- LEÓN, Margarita; GUALDRÓN, Nelsy. **Mejoramiento de la cosecha y poscosecha del banano criollo musa sapientum l. De los productores asociados a san vicente de chucuri; Santander, Colombia.** 2010, 555.
- LI, Jie; CHU, Chao-Hsien; WANG, Yunfeng; YAN, Weili. **An improved fuzzy c-means algorithm for manufacturing cell formation.** 2002. En 2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291) (Vol. 2, pp. 1505-1510). IEEE. doi:10.1109/FUZZ.2002.1006729
- McAULEY, John. **Machine Grouping for Efficient Production.** 1972, (February).
- MEJÍA, Rafael. **Balance preliminar de 2015 y perspectivas de 2016 - SAC - Sociedad de Agricultores de Colombia. 2016.** Recuperado el 4 de febrero de 2017 de: <http://www.sac.org.co/es/estudios-economicos/balance-sector-agropecuario-colombiano/290-balance-y-perspectivas-del-sector-agropecuario-2012-2013.html>
- MinAgricultura; DANE. **El cultivo de la mora de Castilla (Rubus glucus Benth) frutal de clima frío moderado, con propiedades curativas para la salud humana.** 2013. Insumos y Factores Asociados a la Producción Agropecuaria, 17, 1-64.
- MOSTACEDO, Bonifacio; FREDERICKSEN, Todd S. . **Manual De Metodos Básicos De Muestreo Y Análisis En Ecología Vegetal.** 2000. Manual de metodos Basicos de Muestreo y Analisis en Ecología Vegetal, 1-92. doi:10.1007/s13398-014-0173-7.2
- NÚCLEO AMBIENTAL S.A.S. **Cámara de Comercio de Bogotá.** 2015, 1-50.
- PIÑEROS, Clara Janneth. **Recopilación de la investigación del sistema productivo papa criolla.** 2009. Fedepapa, 152.
- SAIZ, Francisco. **Experiencias en el Uso de Criterios de Similitud en el Estudio de Comunidades* Experiences in the use of similarity criteria in community studies.** 1980. Arch. Biol. Med. Exp, 13, 387-402.
- SIPSA. **Cultivo del Limón o Lima Tahití (Citrus latifolia Tanaka) Frente a Los Efectos De Las Condiciones Climáticas Adversas.** 2015. Boletín Mensual Insumos y Factores Asociados a La Producción Agropecuaria. Recuperado de: <https://www.dane.gov.co/index.php/estadisticas-por-tema/agropecuario/sistema-de-informacion->

de-precios-sipsa/componente-insumos-1/componente-insumos-historicos?phpMyAdmin=3om27vamm65hhkhrtgc8rrn2g4

SOKAL, Robert R. . A Quantitative Approach to a Problem in Classification Author(s): Charles D. Michener and Robert R. Sokal Source: 2015, 11(2), 130-162.

WILCHES, Orlando. Métodos Estadísticos Multivariados en Biología Molecular y su Aplicación en Investigación Agrícola. 1995, (1), 66-71.

YIN, Yong; YASUDA, Kazuhiko. Similarity coefficient methods applied to the cell formation problem: A taxonomy and review. 2006. International Journal of Production Economics, 101(2), 329-352. doi:10.1016/j.ijpe.2005.01.014