

1                   **Procedimientos de Escalamiento Multidimensional y Mapas Auto**  
2                   **Organizados para visualizar información bibliométrica relacionada con**  
3                   **Métodos Estadísticos no paramétricos**

6                   **Magaly Herrera Villafranca<sup>1</sup>, Caridad W. Guerra Bustillo<sup>2</sup>, Verena Torres**  
7                   **Cárdenas<sup>1</sup>, Yeter Caraballo Pérez<sup>1</sup>**

9                   <sup>1</sup> *Instituto de Ciencia Animal, Km 47½ Carretera Central, San José de las Lajas.*

10                  <sup>2</sup> *Universidad Agraria de La Habana, Autopista Nacional y Carr. De Tapaste San José de*  
11                  *Las Lajas, La Habana*

13                  **Introducción**

15                  En la actualidad existe un gran cúmulo de información que aparecen en libros, revistas,  
16                  bases de datos etc., y que serviría como base para dar origen a otras investigaciones, sin  
17                  embargo al encontrarse tan dispersa, resulta muy engorrosa la búsqueda e identificación  
18                  de diferentes temas que se desea investigar según el interés del investigador. En este  
19                  sentido se han desarrolla herramientas de análisis que agrupan estas informaciones por  
20                  temática y otras que permiten visualizar este conocimiento. Estas herramientas se  
21                  conocen con el nombre de Escalamiento Multidimensional (EMD) y Mapas Auto  
22                  Organizados.

24                  La técnica de Escalamiento Multidimensional, parte de la matriz de disimilitudes entre  
25                  objetos, para construir un mapa de la menor dimensión posible, que reproduzca lo más  
26                  fielmente posible las disimilitudes originales (Cuesta, 2005)

28                  Las primeras ideas para la aplicación de esta técnica surgieron a mitad del pasado siglo  
29                  con un trabajo realizado por Torgerson en 1952, quien introdujo las primeras ideas sobre  
30                  el tema (Linares, 2001).

32                  Los Mapas Auto organizados han sido de utilidad para descubrir y profundizar en el  
33                  conocimiento de información. Se han empleado en el análisis bibliométrico de grandes  
34                  volúmenes de información y contenido en las bases de textos, además han sido de gran  
35                  utilidad en estudios médicos para el mapeo de informaciones cerebrales.

37                  Miret (2005) plantea que dentro de las Redes Neuronales Artificiales, los Mapas Auto-  
38                  Organizados de Kohonen (SOM del inglés Self Organizing Maps) se caracterizan por su  
39                  capacidad para clasificar un conjunto complejo de patrones de manera no supervisada.

41                  El objetivo de esta investigación fue aplicar las técnicas de Escalamiento Multidimensional  
42                  y los Mapas Auto Organizados para analizar el empleo de los métodos estadísticos no

1 paramétricos en investigaciones científicas publicadas en revistas agropecuarias, en el  
2 período del 2000 al 2007.

3

4 **Materiales y métodos**

5

6 Se revisaron 1418 artículos, de revistas científicas pertenecientes al fondo bibliotecario  
7 del Instituto de Ciencia Animal, las revistas revisadas en el período 2000-2007 fueron:  
8 Nacionales:

- 9     • Revista Cubana de Ciencia Agrícola (RCCA).  
10    • Revista Ciencias Biológicas (CENIC).  
11    • Revista de Pastos y Forrajes (EPFIH).

12 No nacionales

- 13    • Agrociencia (México)  
14    • Veterinaria de México (México)  
15    • Archivos de Zootecnia. (España)  
16    • Ciencias Pecuarias (Colombia)

17  
18 Los resultados fueron tabulados en una base de datos analizándose las variables: nombre  
19 de la revista, número de los artículos por revista, nombres de los autores, tipo de variables  
20 analizadas (cuantitativas y cualitativas), métodos estadísticos empleados y temáticas  
21 tratadas.

22

23 Para procesar la información por revista se empleó el método de Escalamiento  
24 Multidimensional PROXSCAL (en inglés Multidimensional Sealing of Proximity) utilizando  
25 el software estadístico SPSS versión 16.0 (2007), con esta técnica se obtuvieron las  
26 medidas de calidad de ajuste para el agrupamiento de las revistas por temática y según  
27 Métodos no paramétricos. Las medidas de la calidad de ajuste utilizadas fueron el  
28 coeficiente de esfuerzo (Stress, y Stress bruto normalizado), el coeficiente de congruencia  
29 de Tucker y la dispersión explicada (DAF, en inglés). Se considera un buen ajuste cuando  
30 el resultado obtenido del Stress se encuentra próximo a cero y los valores del DAF y el  
31 coeficiente de congruencia de Tucker son de 0,8 en adelante.

32

33 Para visualizar la información por revista se utilizó el software Viscovery SOMine, (2008),  
34 y el método SOM-Wald Clusters, que combina la ordenación de información local del  
35 mapa con el algoritmo de Cluster Jerárquico Clásico de Ward. (Ward, 1963). Este método  
36 permite agrupar los elementos más representativos de lo general a lo particular y une  
37 aquellos con valores similares en cuanto a cercanía hasta formar un grupo llamado  
38 cluster, lo que es representado en un mapa.

1 Del análisis se obtiene dos mapas, en estos se observan las temáticas y los métodos  
2 estadísticos no paramétricos usados en las revistas estudiadas. La información se  
3 organiza por cluster donde la cantidad de ellos varía según los datos del estudio. Para un  
4 mejor entendimiento se ofrece en la parte inferior una leyenda que explica las  
5 abreviaturas usadas y una escala de matices de colores, del color más intenso al menos  
6 intenso, que muestra de derecha a izquierda, en orden decreciente las frecuencias por  
7 temática.

8 En el procesamiento de los datos se trabajó con los valores originales, debido a las  
9 características que deben tener los datos de entrada para que sean procesados por el  
10 software.

11 **Resultados y discusión**

12 La tabla 1 y 2 y la figura 1 muestran los resultados obtenidos, a partir de la aplicación del  
13 método de Escalamiento Multidimensional (PROXSCAL).

14 En la tabla 1 se observan los resultados de la calidad del ajuste, dado por el Stress bruto  
15 normalizado y el S Stress, que son bajos, por lo que se consideran un ajuste adecuado de  
16 los datos.

17 Tabla 1. Resultado de calidad de ajuste para las temáticas por revistas.  
18

Medida	Resultado
Stress bruto normalizado	0,01539
S Stress	0,04395
Dispersión explicada (D.A.F.)	0,98461
Coeficiente de congruencia de Tucker	0,99228

19 Mora y González (2009) indican que en el Escalamiento Multidimensional, el método  
20 PROXSCAL minimiza el Stress bruto normalizado, estos valores se obtienen mediante la  
21 fórmula introducida por Kruskal en 1964, quien ofreció una guía para su interpretación.

22 Otros de los criterios para validar los resultados obtenidos son el D.A.F. y el coeficiente de  
23 congruencia de Tucker los que se encuentran muy cercanos al 100%, dando validez a la  
24 configuración obtenida con esta información. Cuesta (2005), indica que se suele  
25 multiplicar por 100 el Stress obtenido, lo que facilita la interpretación en términos de  
26 porcentaje.

27 En la tabla 2, se ubican en dos dimensiones los pesos o ponderaciones de las revistas  
28 con las temáticas de aplicación, en esta tabla se muestra la distancia que existe entre las  
29 revistas analizadas por diferentes temáticas, al indicar el grado de similitud entre éstas, de  
30 esta forma, las más similares son aquellas que presentan sus valores más próximos. Los

1 valores pequeños muestran los individuos que más se parecen, mientras que los valores  
2 grandes señalan diferencias entre estos. Entre las revistas que más sobresalen por  
3 dimensiones son, por la uno la RCCA y por la dos Ciencias Biológicas y Veterinaria de  
4 México con los valores superiores positivos y negativos respectivamente.

5  
6 Tabla 2. Ubicación en dos dimensiones de las revistas por temáticas

Revistas	Dimensión	
	1	2
Agrociencia	-.015	.025
Archivo de Zootecnia	.001	-.226
Ciencias Biológicas	-.492	<b>.856</b>
Ciencias Pecuarias	-.199	-.034
Pastos y Forraje	.316	.188
RCCA	<b>.939</b>	-.061
Veterinaria	-.550	<b>-.750</b>

21                    Según,

22 Guerrero y Ramírez (2002) el Escalamiento Multidimensional está basado en la  
23 comparación de objetos o de estímulos, de forma que sí un individuo juzga a los objetos A  
24 y B como los más similares, entonces las técnicas de Escalamiento Multidimensional  
25 colocarán a los objetos A y B en el gráfico de forma que la distancia entre ellos sea más  
26 pequeña que la distancia entre cualquier otro par de objetos.

27  
28 Los resultados obtenidos a partir de la tabla 2 se pueden graficar, estos se muestran en la  
29 figura 1. donde se forma cuatro grupos, dos en ambas dimensiones, en la primera se  
30 agrupan las revistas que presentan un comportamiento similar entre las que se  
31 encuentran, Agrociencia, Archivo de Zootecnia, Ciencias Pecuarias y Pastos y Forrajes  
32 mostrándose sus valores muy cercanos entre si, esta misma dimensión se observar que  
33 la revista RCCA se aleja de las demás con un valor diferente a los anteriores, en la  
34 segunda dimensión se encuentran las revistas Ciencias Biológicas y Veterinaria de  
35 México respectivamente, las cuales forman los otros dos grupos.

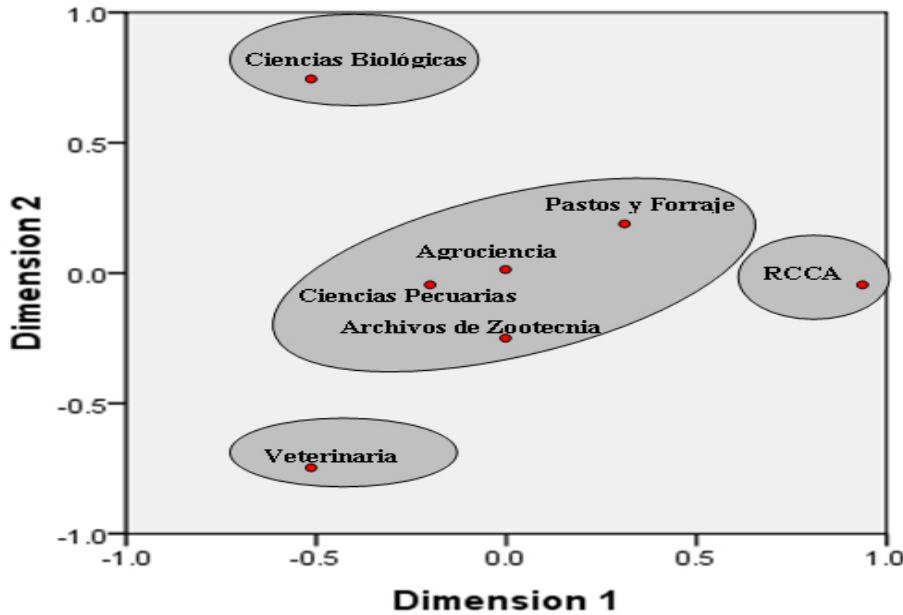
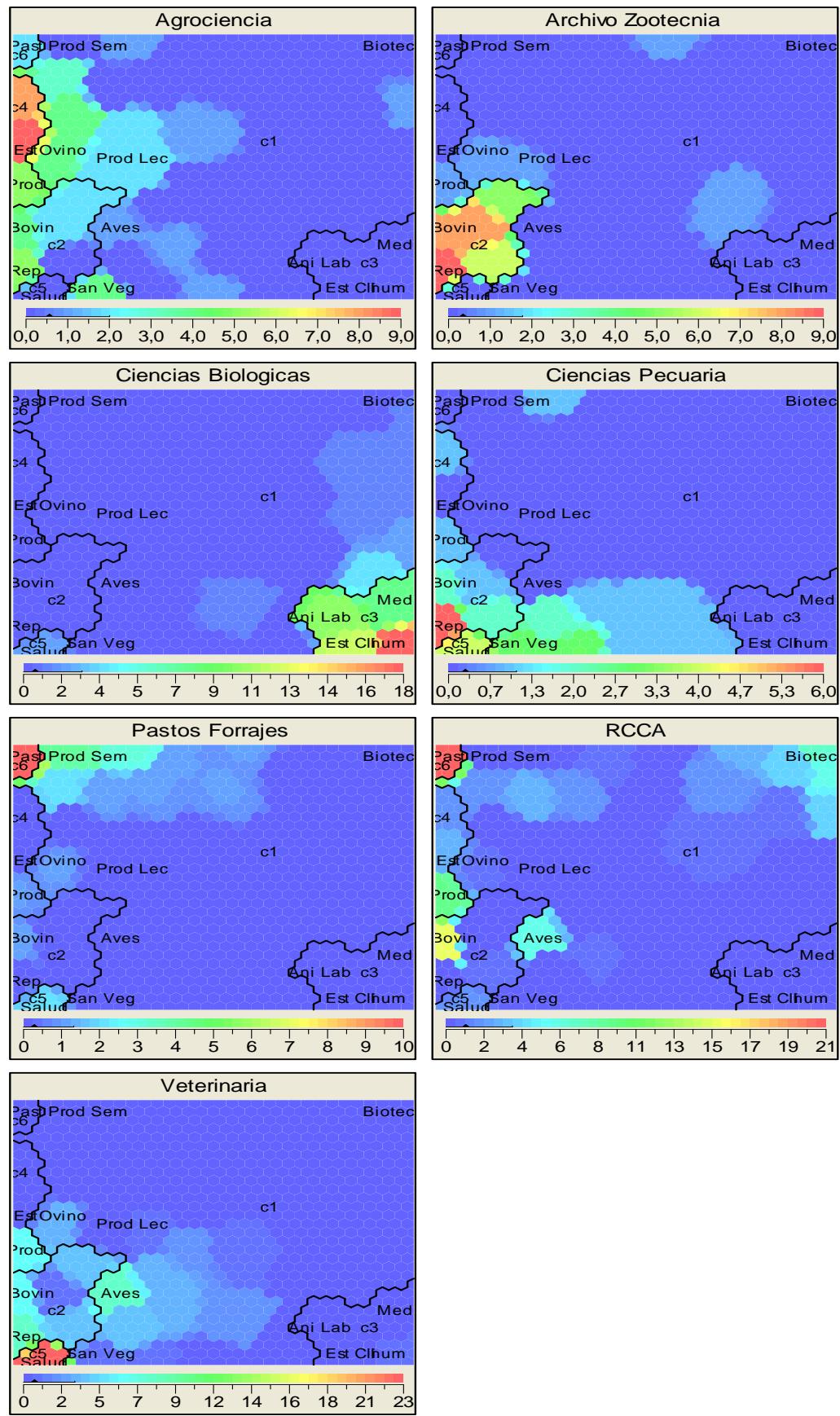


Figura 1. Gráfico en dos dimensiones de las Revistas por temáticas.

En el figura 2 se visualizan las temáticas más abordadas en el período analizado, a través de la técnica de los Mapas Auto Organizados. Esta permitió identificar aquellas temáticas más predominantes en las revistas analizadas. En la Agrociencia se destaca la temática relacionada con los métodos estadísticos aplicados a la Ciencias agropecuarias, en el caso de la revista Archivos de Zootécnica y Ciencia Pecuaria enmarcan sus estudios al tema de Reproducción animal, Ciencias Biológicas a pruebas de medicamentos en humanos. En Pastos y Forrajes y RCCA los temas más estudiados se relacionan con los patos y forraje, haciendo un mayor estudio de la temática la RCCA, y la Veterinaria de México se enfoca a la salud animal.



1  
2

3

Figura 2. Mapa de temáticas más estudiadas por revistas en el período 2000 - 2007

1 Leyenda. Temáticas más representativas  
2 Cluster 1 ( $c_1$ ): investigaciones sobre producción de semillas (Prod Sem), Biotecnología  
3 (Bistec), Ovino, Producción de leche (Prod Lec), Aves, sanidad vegetal (San Veg).  
4 Cluster 2 ( $c_2$ ): estudio en Bovinos (Bovin) y reproducción animal (Rep).  
5 Cluster 3 ( $c_3$ ): estudios médicos (Med), en animales de laboratorios (Ani Lab) y clínicos en  
6 humanos (Est clín hum).  
7 Cluster 4 ( $C_4$ ): métodos estadísticos aplicados a las Ciencias Agropecuarias (Est) y  
8 producción de alimentos (Prod).  
9 Cluster 5 ( $C_5$ ): estudios en Salud Animal (Salud).  
10 Cluster 6 ( $C_6$ ): investigaciones relacionadas con los Pastos (Pas).

11  
12 En la tabla 3 se presentan las medidas de la calidad del ajuste, con la aplicación de los  
13 métodos estadísticos no paramétricos en las revistas analizadas. Se consideran los  
14 mismos indicadores de Stress, así como los otros dos coeficientes de que ponderan la  
15 calidad de ajuste de la configuración dada, la cual resulta superior a la anterior para los  
16 indicadores.

17  
18 Tabla 3. Resultado de la calidad del ajuste de los métodos estadísticos no paramétricos  
19 por  
20 revistas  
21

Medida	Resultado
Stress bruto normalizado	0,00092
S Stress	0,00076
Dispersión explicada (D.A.F.)	0,99908
Coeficiente de congruencia de Tucker	0,99954

22  
23 En la tabla 4, se observa las revistas que tienen una mayor aplicación de los métodos  
24 estadísticos no paramétrica, donde se destacan las Revistas Ciencias Biológicas y  
25 Veterinaria de México, en la primera y segunda dimensión respectivamente.

26  
27 Tabla 4 Ubicación en dos dimensiones de las revistas por temáticas que con mayor  
28 frecuencia abordan el empleo de los métodos estadísticos no paramétricos  
29

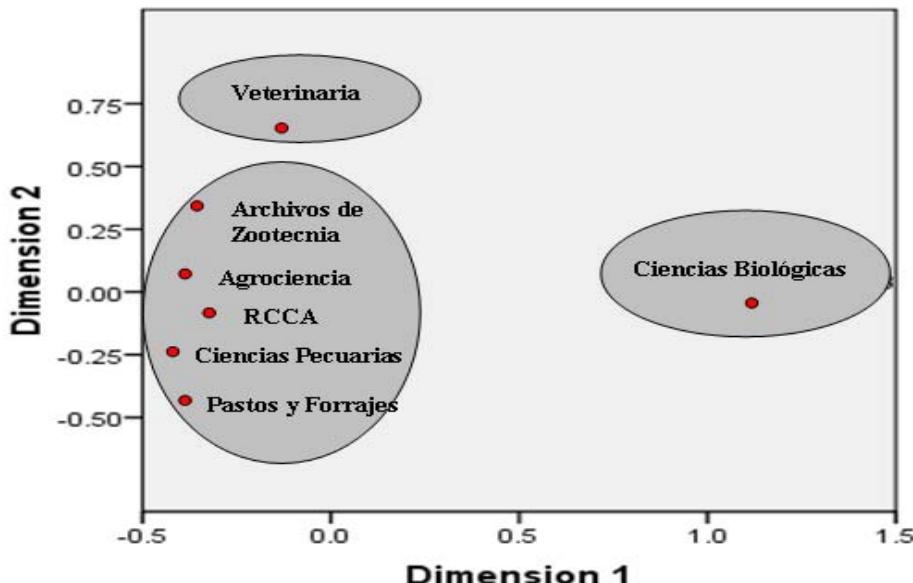
Revistas	Dimensión	
	1	2
Agrociencia	-.302	.109
Archivos de Zootecnia	-.120	.169
Ciencias Biológicas	<b>1.368</b>	-.047
Ciencias Pecuarias	-.232	-.311

Pastos y Forrajes	-.193	-.463
RCCA	-.396	-.088
Veterinaria	-.125	<b>.629</b>

1

2 En la figura 3, se aprecia la formación de tres grupos, dos de estos sobresalen de manera  
 3 independientes con las revistas indicadas en las dimensiones 1 y 2 respectivamente, y el  
 4 tercero se forma con el resto de las revistas que tienen un comportamiento semejante.

5

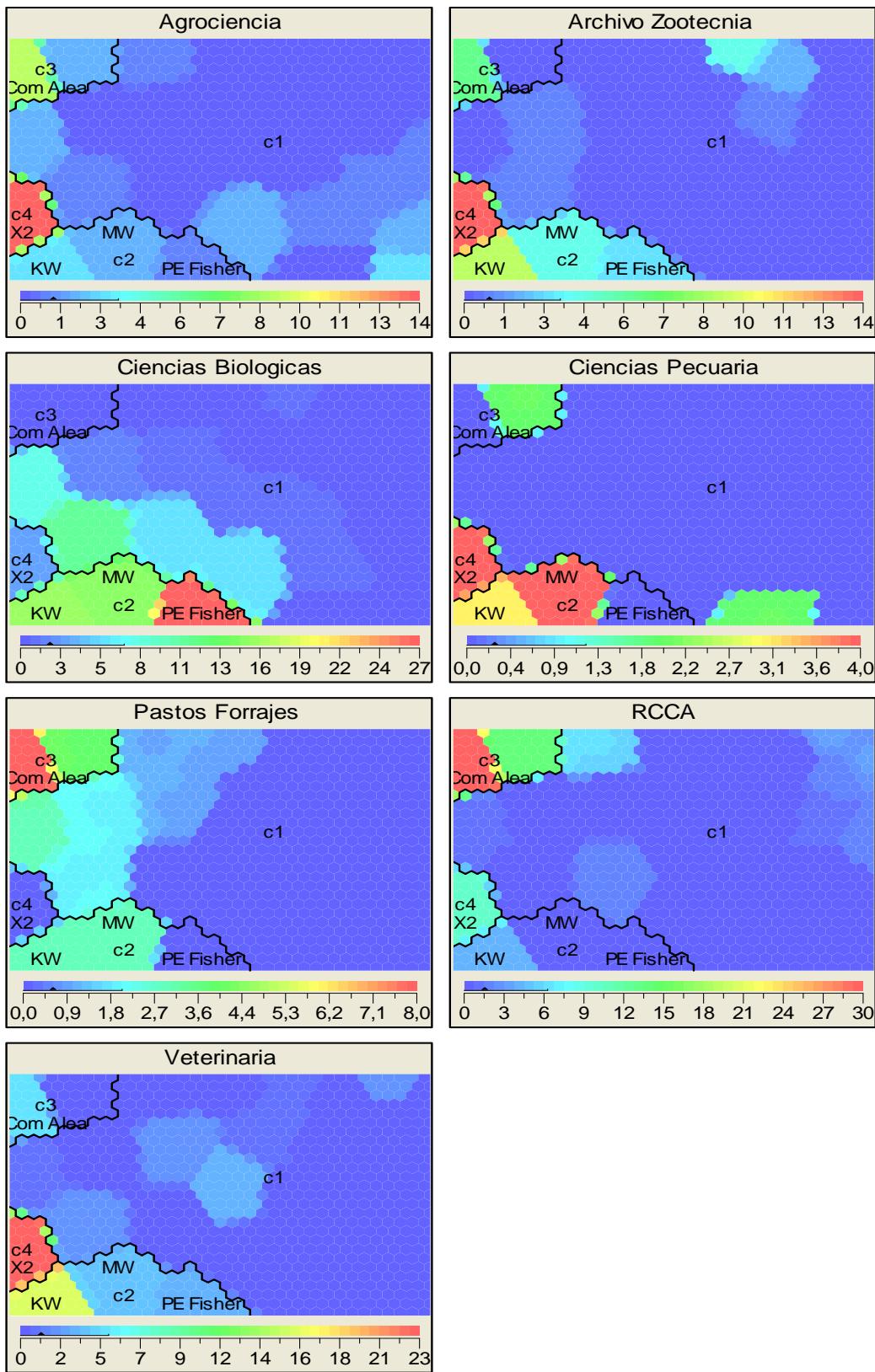


6

7 Figura 3. Gráfico en dos dimensiones de las revistas por temáticas de Estadísticas no  
 8 Paramétricas.

9

10 En la figura 4, se reflejan los mapas referentes a la aplicación de los Métodos  
 11 estadísticos no paramétricos, en este sentido sobresale la Revista Ciencias Biológicas  
 12 con las dócimas Prueba Exacta de Fisher, seguida por Mann-Whitney y Kruskal-Wallis, y  
 13 en la Veterinaria la dócima  $\chi^2$ .



1  
2 Figura 4. Mapa de Métodos Estadísticos no paramétricos más usado en estudios  
3 agropecuarios en el período (2000-2007)  
4  
5  
6

1 Leyenda. Métodos Estadísticos no paramétricos más representativas  
2 Cluster 1 ( $c_1$ ): modelos mixtos, dócimas Kolmogorov-Smirnov y Wilconxo  
3 Cluster 2 ( $c_2$ ): Diseño no paramétrico Completamente Aleatorizado Kruskal Wallis (KW),  
4 Mann Whitney (MW) y Prueba Exacta de Fisher (PE Fisher)  
5 Cluster 3 ( $c_3$ ): Diseño paramétrico Completamente Aleatorizado (Com Alea) y Diseño  
6 paramétrico Bloque al Azar.  
7 Cluster 4 ( $C_4$ ): Dócima de  $X^2$ .

8  
9 Un estudio similar al realizado en este trabajos fue publicado por De Calzadilla (1999), en  
10 el período 1987-1997, encontrando 65 aplicaciones referidas a estos métodos, en 2291  
11 artículos, en 11 revistas científicas, 6 nacionales y 5 extranjeras. Esta autora encontró  
12 que los métodos estadísticos no paramétricos más empleadas fueron, la prueba de  
13 Bondad de ajuste, la correlación por rango Spearman, la prueba Chi-cuadrado de  
14 independencia, la prueba Kruskal-Wallis, Mann-Whitney respectivamente, que coinciden  
15 en cierta medida con los reportado para las revistas Ciencias Biológicas y Veterinaria de  
16 México

17  
18 En el análisis se evidencia que la prueba de bondad de ajuste no ha sido muy empleada,  
19 sin embargo Siegel y Castellan (1995) y Salcines (2003) plantean que resulta más potente  
20 cuando las muestras son más pequeñas.

21  
22 Uno de los métodos más empleado, es el Análisis de Varianza por rango de Kruskal  
23 Wallis, Balzarini et al. (2001), plantean que este procedimiento fue propuesto por Kruskal  
24 Wallis en 1952 y permite por una parte, comparar las esperanzas de dos o más  
25 distribuciones sin la necesidad de verificar el supuesto de que los términos de error se  
26 distribuyen normalmente.

27  
28 En la RCCA, se destaca el trabajo de Mazorra, (2006), donde se hace referencia al  
29 empleo de la dócima de Kruskall-Wallis, porque las variables “cantidad de ramoneo” y  
30 “tiempo de ramoneo” incumplen con el supuesto de homogeneidad de varianza. En este  
31 artículo el autor emplea la dócima de rangos múltiples propuesta por Siegel y Castellan  
32 (1995), con el objetivo de detectar diferencias significativas entre los tratamientos.

33  
34 En la Revista Archivos de Zootecnia, sobresale el trabajo de Delgado et al. (2000), que  
35 refiere su estudio a variables nominales (color de la mucosa, forma de la oreja, etc.) y el de  
36 Revalti, et al. (2004) en un estudio similar, aplicando el Análisis de Varianza Kuskal-  
37 Wallis.

38

1 Relacionados con la dócima Mann-Whitney se encuentran los trabajos de Carvajal et al.  
2 (2001) y Beiro et al. (2001), de la Revista Ciencias Biológica, los cuales estudian las  
3 variables “tiempo de supervivencia” y “actividad antitumoral”, después de aplicados los  
4 tratamientos.

5  
6 Resulta importante destacar el empleo de la dócima de la Probabilidad exacta de Fisher,  
7 la cual es sumamente útil para analizar datos discretos, nominales u ordinales, cuando las  
8 dos muestras independientes son pequeñas. En este caso se reportan los estudios  
9 realizados en la Revista Ciencias Biológicas, Bacardí et al (2004), que evalúa el grado de  
10 toxicidad de un estimulador de colonias de granulositos, y Gámez et al. (2005), que  
11 estudia las “frecuencias de muertes, aparición de signos y lesiones cuando es aplicado el  
12 tratamiento D-003”, en este sentido se observa que las variables analizadas están en  
13 escalas nominal y ordinal.

14  
15 Aunque no fue objetivo de este trabajo analizar el empleo de los Métodos Estadísticos  
16 paramétricos, se puede indicar que en la RCCA y Pastos y Forrajes se hace un mayor  
17 empleo de estos, en lo fundamental en los Análisis de Varianza asociado a los Diseños  
18 Completamente Aleatorizado y Bloque al Azar.

19  
20 Se puede concluir que los métodos estadísticos no paramétricas más empleados, fueron  
21 las dócimas  $\chi^2$ , Mann-Whitney, Kruskal Wallis, y Exacta de Fisher, aunque se observa  
22 más aplicaciones de los métodos estadísticos no paramétricos en investigaciones  
23 científicas publicadas en Revistas Agropecuarias, estos son pocos diversos. El uso del  
24 Escalamiento Multidimensional y los Mapas Auto Organizados (SOM), resultaron muy  
25 útiles para el procesamiento, visualización y análisis de la información, al permitir una  
26 adecuada identificación de las revistas y las temáticas analizadas y el empleo de los  
27 métodos estadísticos no paramétricos en las Ciencias Agropecuarias. Se sugiere  
28 continuar investigando en Revistas de alto índice de impacto, de las diversas ramas  
29 investigativas del Instituto de Ciencia Animal, el empleo de los métodos estadísticos no  
30 paramétricos y ampliar el uso de las herramientas de Escalamiento Multidimensional y los  
31 Mapas Auto Organizados, por su utilidad en la visualización y análisis de información, en  
32 particular para variables categóricas.

33 **Bibliografía**

34  
35 Bacardí, D., Moreno, N., Cosme, K., Porras; Carrera, I., Díaz, G.; Suárez, J.; Aldana, L.,  
36 Amaya, R., 2004. Evaluación de Toxicidad subcrónica Del factor estimulador de  
37 colonias de granulositos. Revista de Ciencias Biológicas 35:3. Cuba  
38

- 1 Balzarini, M. G., Casanoves, F.; Di Rienzo, J. A.; González, L. A., Robledo, C. W. 2001:  
2 Manual del Usuario Insostat. Argentina.
- 3
- 4 Beiro, O., Hernández, D., Arúz, L.; Meriño, N., Álvarez X. y Quintero, Y. 2001. Subnitrato  
5 de bismuto en la actividad antitumoral y la nefrototoxicidad por Cisplanio en ratones  
6 trasplantados con leucemia P-388. Revista de Ciencias Biológicas 32:2 pp. 103.
- 7
- 8 Carvajal, D., Arruzazabala, L., Molina, V.,y Valdés S. 2001. Efecto de D-002 en modelos  
9 experimentales de anafilaxia. Revista Ciencias Biológicas 32:2 pp.93. Cuba.
- 10
- 11 Cuesta, A., 2005. Análisis Multivariante. Dpto. Matemática, Estadística y Computación.  
12 Universidad de Cantabria España 55 p.
- 13
- 14 De Calzadilla, J.1999. Procedimientos de la Estadística no paramétrica. Aplicaciones  
15 en las Ciencias Agropecuarias. Tesis presentada en opción al título de Master en  
16 Matemática Aplicada a las Ciencias Agropecuaria.
- 17
- 18 Delgado, J., Dieguez, E.; Cañuelo, P. 2000. Caracterización exteriorista de las variedades  
19 de cerdo ibérico basadas en característica cualitativas 49:185-186. Revista Archivos  
20 de Zootecnia pp. 201-207. España
- 21
- 22 Gámez, R.; Más, R., Noa, M.; y García, H.; Felip, E; Goicochea, E, Gómez, A.;  
23 Rodríguez; M. y González, J. 2005. Carcinogénesis del D-003 en ratas Sprague  
24 Dawley. Reporte de 1 año. Revista Ciencias Biológicas: 36:3 pp. 201. Cuba
- 25
- 26 Guerra, C. W., De Calzadilla, J. y Torres, V. 2000. Índice de eficiencia en relación con  
27 procedimientos de la estadística no paramétrica. Revista Cubana de Ciencia Agrícola  
28 34:1 pp.1-4. Cuba
- 29
- 30 Guerrero, F. M., Ramírez, J. M. 2002. El análisis de escalamiento multidimensional: una  
31 alternativa y un complemento a otras técnicas multivariantes. Asociación Española de  
32 Profesores Universitarios de Matemática para la Economía y la Empresas. X Jornada  
33 2002
- 34
- 35 Kohonen, T. Self-organizing maps. Berlin: Springer, 3. ed; 2001.
- 36
- 37 Linares, G. 2001. Escalamiento multidimensional. Conceptos y enfoques. Revista  
38 Investigación Operacional. Vol. 22. No. 2. P. 173\_183.
- 39
- 40 Mazorra, C.; Borge, G.; Blanco, M., Borroto, A.; Sorís, A., L 2006. Influencia de la dosis de  
41 cloruro de litio en la conducta de ovinos condicionados que pastorean en plantaciones  
42 de cítricos. Revista Cubana de Ciencia Agrícola 40:4 pp. 425-431. Cuba.
- 43
- 44 Mederos, M.V.; Linares, G.; Miret, E. 2004. Una comparación de dos métodos de  
45 gradiente en el Escalamiento Multidimensional. Revista Ciencias Matemáticas. Vol. 22.  
46 No. 1. p. 44-56.
- 47
- 48 Miret, E. 2005. Un enfoque unificado para técnicas de representación euclidianas. Tesis  
49 para optar por el grado de Doctor en Ciencias Matemática. UH. 148 p.
- 50
- 51 Mora, M.C.; González, M.I. 2009. Caracterización del fracaso empresarial en Colombia:  
52 Un enfoque con Escalado Multidimensional. Universidad de Salamanca. España.  
53 Documento de Trabajo 11/09. 37p.
- 54

- 1 Revialti, M. A.; Capellari, A.; Delgado, J. V. 2005. Recurso genético porcino autóctono en  
2 el nordeste de la republica de Argentina. Revista Archivos de Zootecnia 54:205.  
3 España. [http://www3.uva.es/empresa/uploads/dt\\_11\\_09.pdf](http://www3.uva.es/empresa/uploads/dt_11_09.pdf)
- 4
- 5 Salcines, J. 2003. La normalidad en el gasto de los consumidores: Un análisis de los  
6 hogares de España. Momento económico 126: p .2-14.
- 7
- 8 Siegel, S., Castellan, N.J. 1995. Estadística no paramétrica aplicada a las Ciencias de la  
9 Conducta. Cuarta edición. Editorial Trillas, México. 437p.
- 10
- 11 Visauta, B. 2000. SPSS para Window versión 16.0.
- 12
- 13 Viscovery Software GmbH. 2008. Viscovery SOMine 5.0.2. [Software de cómputo].  
14 Viscovery Software GmbH.
- 15
- 16 Ward, J. 1963. Hierrarchical grouping to optimize an objective function. Jornal of the  
17 Amarican Statistical Association. Vol. 58. No. 4 p. 236\_244.