

Método de estratificación de territorios basado en sistemas de información geográfica y medidas de similitud geométrica
Territory Stratification Method based on Geographic Information Systems and Geometric Similarity Measures

Liset González Polanco ^{1*}	0000-0001-9361-6856
Yadian Guillermo Pérez Betancourt ¹	0000-0003-0354-5586
Roxana Cañizarez González ¹	0000-0002-6143-0616
Liliet de la Caridad González Polanco ¹	0000-0003-4237-5989

¹ Universidad de las Ciencias Informáticas (UCI). La Habana, Cuba.

RESUMEN

Una meta del sistema de salud es la prevención de enfermedades, por ello cobra especial importancia el estudio de la relación de enfermedades con el espacio. Existen evidencias del empleo de los Sistemas de Información Geográfica en estudios sobre la distribución espacial de problemas de salud. A pesar de esto, los trabajos reportados en la literatura consultada no explotan la componente espacial de los datos, lo que limita su integralidad. Por otra parte, existe dispersión en las metodologías, herramientas y técnicas para abordar estudios de este tipo. En esta investigación se presenta un método de estratificación de territorios basado en Sistemas de Información Geográfica y medidas de similitud geométrica, definidas a partir de los criterios: distancia, tamaño y conectividad.

La propuesta permite realizar estudios estratificados según la primera ley de la geografía y garantiza la obtención de estratos más compactos. El método propuesto cuenta con cinco etapas: Selección de indicadores y territorios, Preprocesamiento de indicadores, Agrupamiento, Postprocesamiento y Visualización, soportado en una solución informática basada en software libre. Como parte de la validación se aplica el método en un caso de estudio y se realiza el análisis de índices de validación que avalan la efectividad y competitividad de la propuesta.

PALABRAS CLAVE: estratificación de territorios; medidas de similitud geométrica; sistemas de información geográfica.



ABSTRACT

A goal of the health system is the prevention of diseases, which is why the study of the relationship of diseases with space is of special importance. There is evidence of the use of Geographic Information Systems in studies on the spatial distribution of health problems. Despite this, the works reported in the consulted literature do not exploit the spatial component of the data, which limits its comprehensiveness. On the other hand, there is dispersion in the methodologies, tools and techniques to approach studies of this type. This research presents a method of stratification of territories based on Geographic Information Systems and geometric similarity measures, defined from the criteria: distance, size and connectivity.

The proposal allows for stratified studies according to the first law of geography and guarantees the obtaining of more compact strata. The proposed method has five stages: Selection of indicators and territories, Pre-processing of indicators, Grouping, Post-processing and Visualization, supported by a computer solution based on free software. As part of the validation, the method is applied in a case study and the analysis of validation indices is carried out that guarantee the effectiveness and competitiveness of the proposal.

KEYWORDS: stratification of territories; geometric similarity measures; Geographic Information Systems.

Recibido: 09/01/2021

Aprobado: 15/04/2021

Introducción

El conocimiento de la distribución geográfica de enfermedades, eventos y la situación de salud permite establecer políticas de salud para su atención y establecer las prioridades necesarias en cuanto a recursos ^{(1), (2)}. En este sentido, la estratificación de territorio denota como una valiosa herramienta para analizar el comportamiento de variables en el espacio ^{(3), (4)}. Es considerada como un procedimiento que permite clasificar objetos en clases homogéneas a partir de analogías o relaciones que se establecen entre sus características ⁽⁵⁻⁷⁾. En estudios salubristas suele denominarse estratificación epidemiológica y es parte del proceso integrado de diagnóstico-intervención-evaluación ^{(8), (9)}.



Este documento está bajo [Licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).

La estratificación epidemiológica utiliza varios enfoques para contribuir a la selección de sitios o zonas con problemas de salud y planificar estrategias de intervención. Dentro de los enfoques utilizados se encuentra la estratificación del riesgo y la del riesgo absoluto, por indicadores ponderados, así como los patrones de distribución de frecuencia de los principales riesgos asociados y las técnicas de análisis multivariado fundamentalmente las basadas en el análisis de grupos o agrupamiento ^{(3), (10), (11)}.

La utilización de indicadores ponderados ha sido ampliamente utilizada en estudios salubristas. Este método utiliza un conjunto de indicadores o factores de riesgos asociados al estudio. Los indicadores son ponderados con pesos obtenidos por criterio de expertos en el campo y se obtienen valores para cada territorio, que son utilizados para construir los grupos a partir de rangos establecidos. La principal limitación de esta estrategia radica en el sesgo que se puede introducir al definir los pesos y los rangos para construir los grupos.

Los patrones de distribución de frecuencia de los principales riesgos asociados se utilizan fundamentalmente cuando no se puede determinar o no están identificados los indicadores del riesgo. Su objetivo es identificar zonas donde determinado factor tiene una mayor frecuencia de aparición y luego acometer acciones de intervención.

Cuando es posible cuantificar los indicadores de riesgo y su influencia sobre el área se ha utilizado la estratificación epidemiológica del riesgo. Esta estrategia permite obtener una evaluación sobre el nivel de reducción del problema si se actúa sobre los factores de riesgo. Su principal limitación es que no permite identificar cambios sobre los grupos pues constituye una fotografía del problema en estudio, además suele ser compleja debido a la necesidad de estudios para determinar el riesgo relativo y el riesgo atribuible poblacional.

La distribución del riesgo absoluto se basa en la utilización de un solo indicador para construir los estratos a partir de rangos de distribución según la tasa de incidencia o característica de la región de estudio. Es muy utilizada para la vigilancia epidemiológica a corto plazo pues permite monitorizar determinados indicadores en áreas priorizadas de forma rápida, práctica y oportuna.

Las técnicas de análisis multivariado permiten realizar estudios más complejos porque pueden incorporarse numerosos factores de riesgos y otras variables asociadas al problema. Desde este enfoque la utilización de las técnicas de agrupamiento han sido ampliamente reportadas en la literatura ^{(12), (13)}. Así, se ha utilizado el agrupamiento para construir las clases o estratos a partir de datos temáticos y posteriormente se representa en la cartografía el resultado ⁽¹⁴⁾. Este tratamiento no permite describir relaciones espaciales sobre objetos y por tanto dificulta la incorporación del espacio en el proceso.

En la literatura consultada se propone incorporar la naturaleza espacial a partir de una transformación sobre objetos puntuales o líneas, en la que luego son tratados como temáticos. Esta propuesta no está de acuerdo con la primera ley de la geografía donde



se establece que los objetos en el espacio están relacionados, pero objetos cercanos están más relacionados entre sí que objetos distantes ⁽¹⁵⁾. La utilización de los Sistemas de Información Geográfica (SIG) en el análisis de la distribución espacial de enfermedades ha aumentado considerablemente, sustentado en las herramientas de análisis existentes que posibilitan resolver problemas asociados a la distribución espacial ⁽¹⁶⁾.

Los SIG son herramientas básicas, en todas las esferas del saber, para la confección de mapas digitales y para los análisis geoespaciales, que van más allá de análisis estadísticos y que tributan a una mejor planificación de infraestructura por ejemplo en: estudio demográfico, análisis de vías de transporte, distribución de recursos, distribución y comportamiento de enfermedades. Su desarrollo y aplicación en diferentes áreas ha brindado la posibilidad de analizar grandes volúmenes de datos espaciales.

En Cuba, los métodos aplicados para la estratificación en su mayoría van orientados al análisis estadístico, sin tener en cuenta la naturaleza espacial de los datos, ni el principio de la primera ley de la geografía. Se evidencia una estructuración común que inicia con análisis estadísticos apoyándose en herramientas por ejemplo: Excel, SPSS y luego se presentan los resultados en mapas temáticos utilizando herramientas SIG, por ejemplo, MapInfo; lo cual reduce la eficiencia del trabajo. Estos elementos influyen en el análisis de la relación espacial de indicadores en diferentes áreas y limitan la identificación de estructuras espaciales asociadas a estos indicadores ⁽¹⁷⁾. Las medidas de similitud empleadas consideran las características con igual importancia y están enfocadas a los datos temáticos. Este trabajo tiene como objetivo desarrollar un método de estratificación de territorios basado en Sistemas de Información Geográfica y medidas de similitud geométrica para obtener grupos más compactos.

Método

El aporte fundamental de esta investigación es un método de estratificación de territorios basado en SIG y medidas de similitud geométrica, como continuidad al desarrollo de herramientas y técnicas para el análisis espacial en estudios salubristas. El método está conformado por cinco etapas que cubren los procedimientos identificados en la literatura para este tipo de estudio ^{(4),(15)}. Las etapas propuestas se basan en el enfoque de análisis de datos geoespaciales y se denominan: Selección de indicadores y territorios, Preprocesamiento de indicadores, Agrupamiento, Postprocesamiento y Visualización.

Se sustenta en los siguientes principios:



Este documento está bajo [Licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).

Integración de medidas de similitud geométricas y SIG para darle tratamiento a la componente espacial de los datos en estudios salubristas y una ficha de diagnóstico que permite generar hipótesis o corroborar hallazgos epidemiológicos.

Modelación de la información en torno a la estratificación, los datos geoespaciales y los estudios salubristas. Reutilización de buenas prácticas relacionadas con estudios estratificados como base para el desarrollo del método y su realización mediante analíticas de datos, que favorezcan la incorporación del espacio en este tipo de estudio.

Los enfoques de la propuesta son:

Holístico con el estudio de los indicadores, el espacio en su conjunto y su complejidad, se identifican interacciones, particularidades y procesos que por lo regular no se perciben si se estudian los indicadores por separado y luego se llevan a la cartografía.

Estratégico con la identificación de la situación de los territorios que se estudian y la creación de la ficha de diagnóstico que facilita el establecimiento de objetivos claros a largo plazo y su conjunto de acciones a corto plazo para dar respuesta a las oportunidades y amenazas que impone el entorno, así como las fortalezas y debilidades.

Las cualidades que distinguen al método:

Integración: el método integra indicadores de salud y medidas de similitud geométricas en la estratificación de territorios para facilitar la incorporación del espacio en estudios salubristas. También se distingue por la integración de técnicas de análisis de datos geoespaciales en una solución informática que sirve de soporte tecnológico.

Usabilidad: el enfoque de guía para la estratificación y la interfaz de la instanciación facilitan la integración de la cartografía e indicadores en los estudios sin necesidad de mucho dominio en este campo.

Fiabilidad: la información que brinda se corresponde con el análisis de los indicadores aportados.

Flexibilidad: a partir del uso de indicadores de naturaleza variada y un marco de trabajo para la estratificación, se facilita adaptarse a cambios que se deseen incluir en los estudios.

Resultados

En la figura 1 se muestra el método propuesto, así como la interacción entre sus etapas y componentes computacionales. Las entradas del método son:



Este documento está bajo [Licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).

Base cartográfica: que está formada por capas vectoriales, de las cuales al menos una debe ser de polígonos y es la capa base que se utilizará como definición de las entidades administrativas a utilizar en la estratificación. El resto de las capas pueden ser de puntos, líneas o polígonos y responden a indicadores geoespaciales.

Fuente estadística: cuenta con los indicadores estadísticos a utilizar en la estratificación y sus mediciones se deben corresponder con las entidades geoespaciales definidas por la capa base para la estratificación.

Las salidas del método sirven de soporte para la toma de decisiones, aportan elementos asociados a distribuciones y procesos espaciales útiles para la definición de objetivos y planes en el tratamiento a problemas de salud. Estas son:

Ficha de diagnóstico: contiene los elementos identificados en la estratificación, constituye una base para la toma de decisiones. En esta se detallan los estratos y territorios más afectados, así como los posibles factores asociados. También dispone de un mapa temático a través del cual se puede interactuar con los estratos y territorios.

Mapa temático: contiene los estratos clasificados a través de una escala de colores que responde al valor del riesgo obtenido.

Con la utilización de este método es posible identificar los territorios y estratos más afectados, de conjunto con los posibles factores asociados, que permite a los salubristas la selección de intervenciones y adecuación de los servicios. También permite la evaluación de los resultados de las intervenciones realizadas, así como el monitoreo y ajuste de indicadores de acuerdo a los problemas detectados.



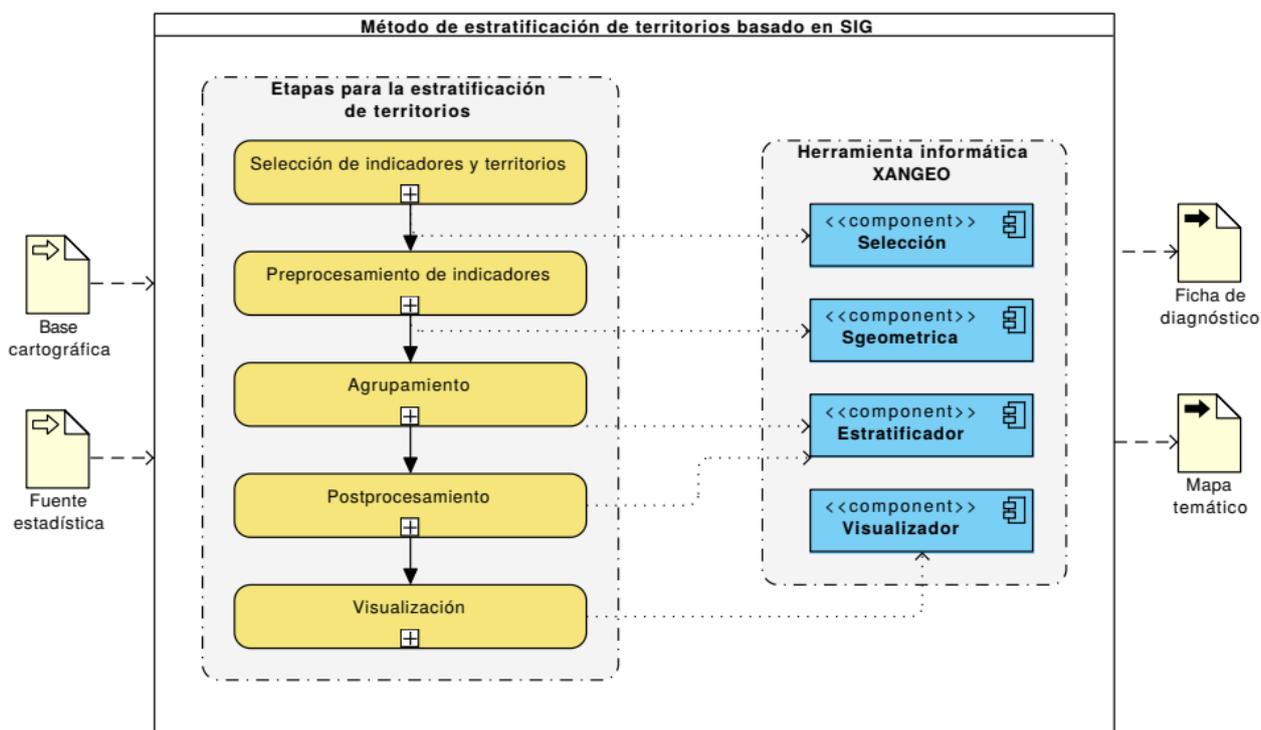


Fig.1- Método de estratificación de territorios basado en SIG y medidas de similitud geométrica.

Descripción de las etapas que conforman el método propuesto

Selección de indicadores y territorios

Esta etapa tiene como objetivo elegir los indicadores que se utilizarán en la estratificación y los territorios que serán objeto de estudio, por lo que es necesario disponer de un mapa vectorial e indicadores disponibles en fuentes estadísticas, o recopilados por el investigador utilizando diferentes técnicas y herramientas. Los indicadores pueden ser de naturaleza espacial o temática. Los datos espaciales provienen de la geografía del terreno, y los datos temáticos corresponden a la información de otras fuentes como son datos de la población, factores de riesgo e indicadores de salud, por solo mencionar algunos.

La selección de indicadores estadísticos parte de la identificación de los indicadores disponibles en la fuente estadística. Los indicadores se escogen de forma manual o de forma automática. La forma manual consiste en que el investigador tiene identificados cuales son los que inciden en el problema objeto de estudio o cuando realiza una evaluación de los resultados de las intervenciones realizadas a partir de estudios realizados, e incluye el monitoreo y ajuste de indicadores de acuerdo a los problemas detectados.



En la selección de atributos de forma automática se reduce la dimensionalidad del conjunto, a través de la selección del subconjunto de mejor desempeño bajo algún criterio de clasificación. La selección de atributos reporta beneficios al eliminar los rasgos irrelevantes y redundantes, lo que posibilita una mejor representación de la información original y una disminución del costo computacional. En la selección de atributos de forma automática están relacionadas distintas áreas como la del reconocimiento de patrones, el aprendizaje automático y la minería de datos.

Los procedimientos de selección de atributos requieren un método de generación de subconjuntos (basado en un proceso de búsqueda). En la presente investigación se utiliza la metaheurística basada en colonia de hormigas (ACO) ⁽¹⁸⁾. Las hormigas construyen las soluciones de manera probabilística, guiándose por un rastro de feromona artificial y por una información calculada a priori de manera heurística. Se obtienen las características más visitadas, es decir el subconjunto de características que mejor describe a los objetos.

Al finalizar el investigador tiene la opción de incluir los indicadores que estime necesarios para su estudio y que no se encuentren en la solución encontrada por el procedimiento de selección. El resultado de esta etapa es un conjunto de indicadores, tanto estadísticos como geoespaciales, que se utilizarán para realizar la estratificación. En la siguiente etapa estos indicadores son preprocesados para facilitar el análisis y la interpretación de los resultados.

Preprocesamiento

Tiene como objetivo preparar la base cartográfica y los indicadores seleccionados para realizar la estratificación. Posteriormente se calcula el aporte informacional y se normalizan los valores de los indicadores seleccionados. La base cartográfica debe estar formada por n capas con $n \geq 1$, para el caso donde $n = 1$ entonces, la capa debe ser de polígonos y contiene los objetos geoespaciales asociados a los territorios objetos de estudio.

Agrupamiento

Esta etapa tiene como objetivo la construcción de los estratos. Para ello se clasifican los territorios en grupos homogéneos (estratos), utilizando técnicas y algoritmos de agrupamiento. Desde el enfoque de la minería de datos geoespaciales destaca la utilización de los algoritmos de agrupamiento en estudios estratificados. Una clasificación general los divide en tres grupos fundamentales: particionales, jerárquicos y basados en densidad o categorías según el procedimiento que utilizan para agrupar los objetos. La elección de dicho algoritmo depende de varios factores, por ejemplo: tipo de datos disponibles, criterio de agrupamiento, complejidad, detección de valores atípicos, el propósito particular y de la experiencia del investigador o del equipo que ejecuta la investigación.



Postprocesamiento

Tiene como objetivo la obtención de los territorios y estratos más afectados por indicadores. Se debe especificar que esta etapa en la estratificación de territorios recibe los datos resultantes del Agrupamiento. Se realiza el análisis de la auto-correlación espacial de los indicadores por estratos. De esta forma se determinan los que aumentan el índice de riesgo.

Visualización

Su objetivo es representar en un mapa temático del tipo coropleto o coroplético cada grupo homogéneo de territorios, esencial para comprender las realidades espaciales. Los mapas temáticos coropletos reportados en la literatura tienen la particularidad de que las regiones se colorean a partir de una medida estadística, como puede ser la densidad de población o el ingreso por habitante. Este tipo de mapa facilita la comparación de una medida estadística de una región con la de otra o muestra la variabilidad de esta para una región dada. Se reporta en fenómenos como la variación espacial que coincide con los límites de las unidades de medición, por ejemplo se distribuye de manera uniforme dentro de cada unidad de medición. En esta investigación cada grupo homogéneo de territorios es representado por un color asignado al estrato en función del riesgo promedio del estrato.

Con el objetivo de demostrar la viabilidad del método propuesto se implementa un complemento para el SIG QGIS. QGIS es una aplicación de escritorio, con un entorno sencillo, amigable, intuitivo y fácil de utilizar ⁽¹⁹⁾. El diseño del sistema y la organización está regido por un estilo de Arquitectura en capas, la cual posibilita que la funcionalidad del sistema este organizada en capas separadas y cada una apoya sólo en las facilidades y los servicios ofrecidos por la capa debajo de ella. Este enfoque en capas soporta el desarrollo incremental del sistema.

Aplicación del método propuesto a un caso de estudio

Para valorar los resultados de la solución propuesta se decide aplicar un caso de estudio, en correspondencia con el trabajo realizado por Companioni ⁽¹⁴⁾, en el cual la autora realiza una estratificación de territorio en Cuba para analizar el comportamiento de indicadores de salud. Utiliza la división político-administrativa del año 1976, en la cual existía una composición de 14 provincias, selecciona cuatro indicadores y un total de 26 variables del Anuario Estadístico del año 2001.

Se escoge como base para la estratificación una capa de polígonos obtenida desde la Infraestructura de datos espaciales de la República de Cuba (IDERC) ⁽²⁰⁾ con la división político-administrativa vigente para el año 2001. Una vez seleccionadas las fuentes de datos para el estudio, se procede a la configuración de los parámetros, la selección de los territorios y los indicadores. En este estudio se analizan todas las provincias y los indicadores, además se clasifican estos últimos en función de su aporte al riesgo



calculado en la etapa de Preprocesamiento. A partir de los territorios y los indicadores seleccionados se obtienen los grupos y se visualizan en un mapa temático, como se muestra en la Figura 2. La solución computacional permite interactuar con los estratos y con los elementos de la ficha de diagnóstico.

Luego de realizados los estudios estratificados utilizando cada una de las funciones de distancia y similitud propuestas (distancia, conectividad y tamaño), se procede a calcular los índices de validación interno para cada estudio. Se incluye también el cálculo para los estudios presentado por Companioni ⁽¹⁴⁾.

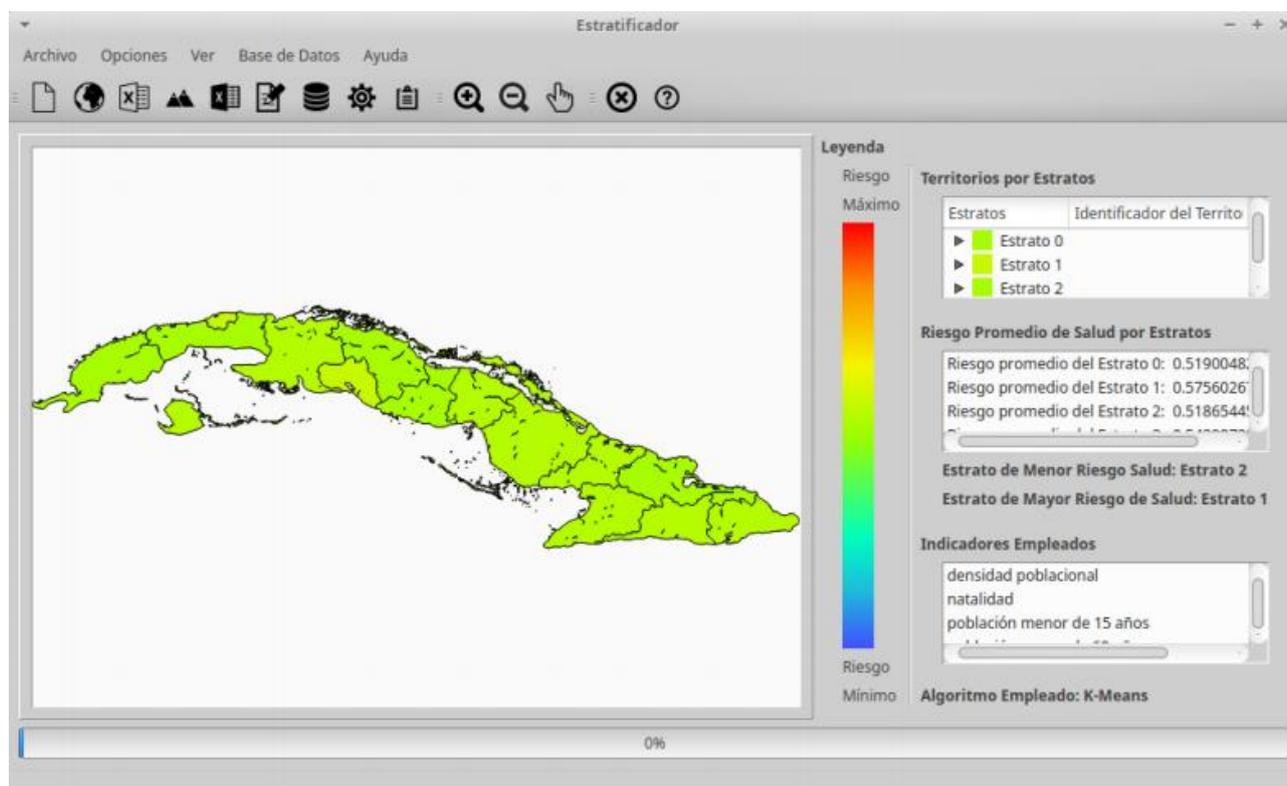


Fig. 2- Visualización de la estratificación.

Discusión

En la Tabla 1 se muestran los valores obtenidos, se evidencia que los resultados utilizando el método propuesto en esta investigación son competitivos y que para el caso del coeficiente de Silhouette que evalúa la calidad del agrupamiento los resultados son superiores.



Tabla 1- Validación con índices internos. Elaboración propia.

Índice	Davies_Bouldin	Calinski_Harabaz	Silhouette_Score
Companioni 2005	1.991	2.182	-0.029
Pérez et Al. 2016	1.748	2.484	0.019
Conectividad	1.810	1.827	0.005
Distancia	1.903	1.199	0.012
Tamaño	1.800	1.247	0.013

Luego se procede a comparar los resultados obtenidos a partir de la utilización de las tres funciones propuestas para evaluar los índices de validación externos precisión (P), Jaccard (J), Fowlkes Mallows (F&M), Rand_Index (Rand), Bray Curtis (BC), V_measure (VM), Mutual_Info (MI), Completeness (C) y Homogeneity (H) .

En la figura 3 se muestran los valores obtenidos teniendo como referencia el trabajo de Companioni ⁽¹⁴⁾ y se evidencia un mejor desempeño de la función conectividad en cuanto a precisión y al índice de Jaccard. Para el resto de índices los mejores resultados de forma general los obtiene la función basada en el tamaño.

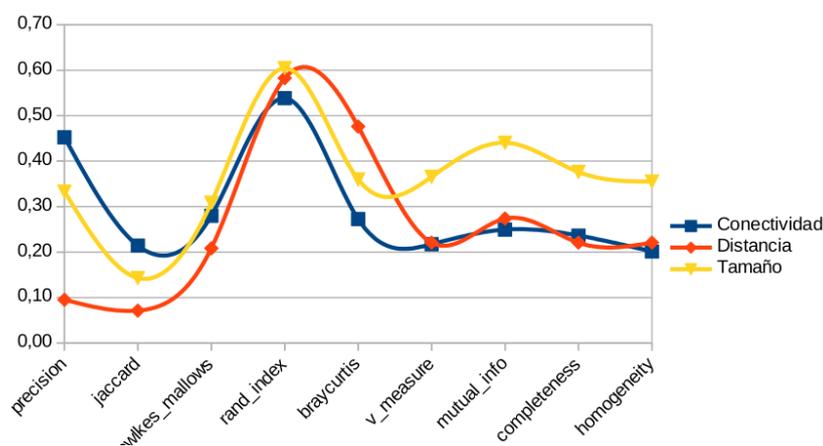


Fig. 3- Resultados de evaluación de las métricas con referencia a Companioni. Elaboración propia.

Se observan mejores resultados en cuanto a precisión y al índice de Jaccard para la conectividad, para el resto, el mejor desempeño lo obtiene la función basada en la distancia entre polígonos. Para todos los casos los mejores resultados se obtienen al evaluar el índice Rand_Index.



Conclusiones

Como resultado de la presente investigación se obtuvo un método para la estratificación de territorio basado en Sistemas de Información Geográfica y medidas de similitud geométrica, para obtener grupos más compactos. En función de los resultados obtenidos se arribó a las siguientes conclusiones:

A partir de la sistematización de los principales referentes teóricos que sustentan la presente investigación, se confirma que las propuestas para la estratificación reportadas en la literatura presentan limitaciones para la incorporación de la componente espacial en el proceso de estratificación de territorios.

La identificación de los constructos dentro de la minería de datos geospaciales facilitó la descripción adecuada del problema y su solución a partir del método propuesto, que además integra los enfoques aportados en investigaciones precedentes con relación a estudios estratificados y la componente espacial de los datos en el espacio de solución del problema.

La instanciación del método propuesto a partir del desarrollo de un componente para el SIG QGIS permitió la evaluación concreta del método a través de la realización de un caso de estudio.

Al evaluar los índices de validación tanto internos como externos se pudo comprobar que si se incorporan medidas de distancia geométricas se obtienen grupos más compactos sin afectar la precisión de la clasificación.

Referencias

1. Rodríguez-Morales AJ, Cárdenas-Giraldo EV, Montoya-Arias CP, Guerrero-Matituy EA, Bedoya-Arias JE, Ramírez-Jaramillo V, et al. Mapping chikungunya fever in municipalities of one coastal department of Colombia (Sucre) using geographic information system (GIS) during 2014 outbreak: Implications for travel advice. *Travel Med Infect Dis.* junio de 2015;13(3):256-8.
2. Patrick SW, Davis MM, Lehmann CU, Cooper WO. Increasing incidence and geographic distribution of neonatal abstinence syndrome: United States 2009 to 2012. *Journal of Perinatology.* agosto de 2015;35(8):650-5.
3. de Araújo Nobre M, Ferro A, Maló P. Adult Patient Risk Stratification Using a Risk Score for Periodontitis. *Journal of clinical medicine.* 2019;8(3):307.
4. Humburg J. Social Stratification and Social Mobility. En: Humburg J, editor. *Television and Precarity: Naturalist Narratives of Poor America* [Internet].



- Stuttgart: J.B. Metzler; 2020 [citado 5 de junio de 2020]. p. 25-31. Disponible en: https://doi.org/10.1007/978-3-476-05660-3_3
5. Delgado Acosta H, González Moreno L, Valdés Gómez M, Hernández Malpica S, Montenegro Calderón T, Rodríguez Buergo D. Estratificación de riesgo de tuberculosis pulmonar en consejos populares del municipio Cienfuegos. *MediSur*. 2015;13(2):275–284.
 6. Montenegro Calderón T, Bermúdez Novoa A, Águila Rodríguez N, Delgado Acosta HM, Suárez del Villar Seuret S. Estratificación de riesgo de tuberculosis pulmonar en consejos populares del municipio Abreus. Cienfuegos, 2015. *MediSur*. agosto de 2019;17(4):486-93.
 7. Pérez Betancourt Y, Betancourt YGP, Polanco LG, Pérez RM, Vega YT. Estratificación de territorios basada en indicadores de salud sobre el Sistema de Información Geográfica QGIS. *Rev Cuba Cienc Inform*. 31 de mayo de 2016;10(0):163-75.
 8. León Cabrera P, Fariñas Reinoso AT, Galindo Reymond K, Prior García A, Dihigo Faz T, Núñez Valdés L. Estratificación epidemiológica del riesgo de las enfermedades emergentes y reemergentes por áreas de salud. Provincia de Matanzas. 2002-2006. *Revista Médica Electrónica*. febrero de 2012;34(1):34-46.
 9. Pérez CG, Aguilar PA. Estratificación epidemiológica de riesgo. *Revista Archivo Médico de Camagüey*. 2013;17(6):762–783.
 10. Quesada Aguilera JA, Quesada Aguilera E, Rodríguez Socarras N. Diferentes enfoques para la estratificación epidemiológica del dengue. *Revista Archivo Médico de Camagüey*. febrero de 2012;16(1):109-23.
 11. Cox J, Sovannaroth S, Soley LD, Ngor P, Mellor S, Roca-Feltrer A. Novel approaches to risk stratification to support malaria elimination: an example from Cambodia. *Malaria journal*. 2014;13(1):371.
 12. Martins ES, Ribeiro M, Lisboa-Filho J, Reinaldo F, Freddo A, Reis LP. Clustering of spatial data for knowledge extraction. En: 2016 11th Iberian Conference on Information Systems and Technologies (CISTI) [Internet]. IEEE; 2016 [citado 3 de agosto de 2016]. p. 1–6. Disponible en: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7521626
 13. Aghabozorgi S, Shirkhorshidi AS, Soltanian TH, Herawan U, Wah TY. Spatial and temporal clustering of air pollution in Malaysia: A review. En: International Conference on Agriculture, Environment and Biological Sciences [Internet]. 2014 [citado 12 de marzo de 2017]. p. 213–219. Disponible en: https://www.researchgate.net/profile/Ali_Seyed_Shirkhorshidi/publication/267395414_Spatial_and_Temporal_Clustering_of_Air_Pollution_in_Malaysia_A_Review/links/544f15a90cf26dda0890270a.pdf
 14. Bombino Companioni Y. Construcción de tipologías: metodología de análisis para la estratificación según indicadores de salud. 2005;
 15. Liu Y. Geographical stratification and the role of the state in access to higher education in contemporary China. *International Journal of Educational Development*. 2015;44:108–117.



16. Malczewski J, Rinner C. GIScience, Spatial Analysis, and Decision Support. En: Multicriteria Decision Analysis in Geographic Information Science [Internet]. Springer; 2015 [citado 9 de febrero de 2016]. p. 3–21. Disponible en: http://link.springer.com/chapter/10.1007/978-3-540-74757-4_1
17. Pérez Betancourt YG, González Polanco L, Febles Rodríguez JP, Cabrera Campos A. Cellular Automata Based Method for Territories Stratification in Geographic Information Systems. En: Botto-Tobar M, León-Acurio J, Díaz Cadena A, Montiel Díaz P, editores. Advances in Emerging Trends and Technologies. Cham: Springer International Publishing; 2020. p. 507-17. (Advances in Intelligent Systems and Computing).
18. Bello R, Nowe A, Gomezd Y, Caballero Y. Using ACO and rough set theory to feature selection. WSEAS Transactions on Information Science and Applications. 2005;2(5):512–517.
19. Flenniken JM, Stuglik S, Iannone BV. Quantum GIS (QGIS): An introduction to a free alternative to more costly GIS platforms. EDIS. 2020;2020(2):7–7.
20. Infraestructura de Datos Espaciales de la República de Cuba (IDERC) [Internet]. [citado 20 de noviembre de 2020]. Disponible en: <http://www.iderc.cu/iderc1/web/iderc/inicio>.

Conflicto de interés

Los autores del trabajo declaramos que no existen conflictos de interés

Declaración de autoría

Liset González Polanco, Máster en Informática Avanzada y Profesora auxiliar. Jefe del departamento de Informática de la Facultad 3 de la Universidad de las Ciencias Informáticas (UCI): fue responsable del diseño de las etapas del método y la introducción de los resultados.

Yadian Guillermo Pérez Betancourt, estudiante de doctorado en Informática, Profesor Auxiliar del Dpto. de Informática de la Facultad 3 de la UCI: desarrolló la solución informática que da soporte al método propuesto y su integración con el Sistema de Información Geográfica.

Roxana Cañizarez González, Doctora en Ciencias, Profesora auxiliar del Dpto. de Informática de la Facultad 3 de la UCI: diseñó los casos de estudio y lideró en la concepción de la propuesta.

Liliet de la Caridad González Polanco, Ingeniera en Ciencias Informáticas, profesora de la Universidad de las Ciencias Informáticas: responsable de la validación e introducción de los resultados.

