



Revista Cubana de Salud Pública

ISSN: 0864-3466

ecimed@infomed.sld.cu

Sociedad Cubana de Administración de Salud  
Cuba

Alegret Rodríguez, Milagros; Grau Abalo, Ricardo; Rodríguez Rodríguez, Mercedes  
La clasificación multivariante de áreas geográficas como una alternativa útil a la investigación  
salubrista

Revista Cubana de Salud Pública, vol. 34, núm. 2, abril-junio, 2008  
Sociedad Cubana de Administración de Salud  
La Habana, Cuba

Disponible en: <http://www.redalyc.org/articulo.oa?id=21419850004>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## La clasificación multivariante de áreas geográficas como una alternativa útil a la investigación salubrista

### Multivariate classification of geographic areas as a useful alternative for public health research

Milagros Alegret Rodríguez<sup>I</sup>; Ricardo Grau Abalo<sup>II</sup>; Mercedes Rodríguez Rodríguez<sup>III</sup>

IDra.C de la Salud. Unidad de Análisis y Tendencias en Salud. CPHE Villa Clara, Cuba.

IIDr.C. Psicológicas. Instituto de Matemática Aplicada. UCLV. Villa Clara, Cuba.

IIIMáster en Ciencias. Instituto Pedagógico Félix Varela. Villa Clara, Cuba.

---

#### RESUMEN

**Introducción** Se ejemplifica el uso de técnicas jerárquicas de aglomeración mediante similitudes multivariadas y su utilidad para correlacionar y explicar estadísticas de daño a la salud, como una herramienta útil a los propósitos de la investigación salubrista involucrando al espacio.

**Objetivos** Identificar constructos multivariantes contextuales en unidades geográficas y relacionarlos con daños a la salud.

**Métodos** Se realizó una clasificación multivariante de los 169 municipios cubanos utilizando técnicas de conglomerados jerárquicos no supervisados. Los ejes analíticos fueron: demográfico, económico y condiciones de vida. Se utilizaron además técnicas de correlación y regresión con indicadores de daño.

**Resultados** Se caracterizaron 4 estratos multivariados y se comprobó el efectivo carácter ordinal que presentan respecto a la respuesta de daño en relación con la morbilidad y mortalidad de las enfermedades crónicas no transmisibles (7 de 11 indicadores), y morbilidad por enfermedades transmisibles (2 de 8 indicadores). Los indicadores del PAMI mostraron independencia de los estratos construidos.

**Conclusiones** La metodología utilizada permite discernir una estructura contextual útil para explicar la distribución de eventos de daño estudiados y aún explicar aquellos que por haber cerrado brechas a través de las acciones de programas, resultan independientes de esa estructura contextual. El tratamiento del espacio

bajo este enfoque es flexible, refractario a los problemas de análisis de áreas pequeñas, contigüidades espaciales y cambios de la unidad espacial de análisis, al crear un macroespacio para fines analíticos ulteriores y estructurar el espacio más allá de las fronteras, estableciendo un marco contextual relativamente homogéneo y estable. Estas herramientas son fundamentales para la planificación y evaluación de intervenciones en salud pública y en general, para la gestión de políticas de salud.

**Palabras clave:** Salud pública, análisis espacial, análisis multivariado.

---

## ABSTRACT

**Introduction** Use of hierarchical clustering techniques based on multivariate similarities and its advantages for correlation and explanation of health deterioration statistics was exemplified, as a useful tool for health research involving space.

**Objectives** To identify multivariate contextual constructs in geographical units and relate them to health damage.

**Methods** A multivariate classification of the 169 municipalities by using unsupervised hierarchical cluster techniques. The analytical axes were demographic, economic and living conditions. Correlation and regression techniques with health damage indicators were also used.

**Results** Four multivariate strata were characterized and their effective ordinal character was confirmed in relation to damage response in terms of morbidity and mortality in non communicable chronic diseases (7 out of 11 indicators) and of morbidity from communicable diseases (2 out of 8 indicators). PAMI indicators showed that they did not depend on the constructed strata.

**Conclusions** This methodology allowed identifying a useful contextual structure to explain the distribution of studied damage events and even those events which are not dependent on this contextual structure, since they have closed gaps by means of planned actions. According to this approach, space is addressed in a flexible way, refractory to problems existing in the analysis of small areas, spatial contiguities and changes in the analytical spatial unit. This is due to the creation of a macro space for further analytical purposes and to the structuring of space beyond frontiers, thus setting a relatively homogeneous and stable contextual frame. These tools are essential to plan and evaluate public health interventions, and for the health policy management in general.

**Key words:** Public health, spatial analysis, multivariate analysis.

---

## INTRODUCCION

La investigación etiológica asociada al espacio no ha sido un terreno fuerte en el campo de la salud, y pocas veces se establece como metodología el contraste de agrupaciones espaciales de casos *versus* riesgo contextual, ya que generalmente

las técnicas de detección de conglomerados asume la premisa de riesgo constante en la población, excepto que se trate de la investigación asociada a una fuente generadora de riesgos.<sup>1</sup> No es usual el abordaje contextual del riesgo y su relación con los resultados de salud, aunque se encuentran buenas contribuciones metodológicas asociadas a las ciencias sociales y ambientales en la obra de *Haining*.<sup>2</sup>

Las técnicas de análisis de conglomerados jerárquicos y no jerárquicos han sido utilizadas con fines de ordenamiento territorial, para el planeamiento urbanístico, el uso de la tierra, la optimización de cultivos y otras aplicaciones,<sup>3,4</sup> muchas de ellas hacen uso de la técnica de clusters para clasificar de forma multivariante a través de atributos relevantes unidades de análisis, sin embargo, se encuentran escasas aplicaciones en el ámbito de los determinantes de salud, a pesar que en el corazón mismo del ordenamiento territorial subyacen aspectos que involucran centralmente todas las relaciones del hombre y su entorno. El valor de estas aproximaciones metodológicas radica en que si el ordenamiento se basa en variables clave a los eventos que desean ser estudiados, este marco contextual resulta una referencia relativamente estable en el tiempo que permite la toma de decisiones, el pronóstico y otras acciones de interés a las ciencias de la salud.

Más allá de las prevenciones necesarias respecto a las posibles falacias resultantes de pretender extrapolar lo individual a lo contextual o viceversa, este enfoque flexibiliza el tratamiento para integrar niveles de análisis grupales e individuales, permite la integración de áreas, exime al espacio de estructuras políticas o delimitaciones administrativas de cualquier tipo; acepta, tras la debida operacionalización, variables provenientes de disímiles ejes y por todo ello se acerca a las exigencias del análisis multinivel. Se puede alcanzar diferentes resultados de clasificación sobre una misma región, pero la calidad de las variables incorporadas, y sobre todo su plausibilidad dentro del rompecabezas etiológico, permitirán alcanzar los mejores resultados.

Este trabajo ejemplifica el uso de técnicas jerárquicas de aglomeración mediante similitudes multivariadas y su utilidad para correlacionar y explicar estadísticas de daño a la salud, como una herramienta útil a los propósitos de la investigación salubrista involucrando al espacio.

## MÉTODOS

Para el desarrollo de este trabajo se utilizó como unidad geográfica el municipio. Las 169 municipalidades cubanas fueron estratificadas en base a tres ejes fundamentales: demográfico, económico y condiciones de vida.<sup>5</sup>

En el eje demográfico fueron medidos los siguientes atributos:

- Estructura por edad de la población.
- Razón de dependencia.
- Escolaridad.
- Urbano/ruralidad.
- Densidad poblacional.

En el eje económico se tuvo en cuenta:

- Actividad económica fundamental.
- Producción mercantil.
- Población laboralmente activa.

El eje condiciones de vida contempló:

- Cobertura de los servicios de acueducto.
- Cobertura de las redes de alcantarillado.
- Recogida de residuales sólidos.
- Calidad de la vivienda.
- Vialidad.
- Presencia de áreas verdes y de recreación, estructuradas.

Las variables fueron analizadas para detectar posibles correlaciones entre ellas mediante matrices de correlación bivariada y finalmente las municipalidades con sus atributos fueron analizadas para conformar conglomerados de similaridades multivariantes mediante técnicas jerárquicas. Se utilizó para ello el programa SPSS v.11.5. Mediante el análisis del historial de aglomeración se determinó el número de estratos de mayor utilidad a los propósitos del análisis, considerando los criterios de máxima cohesión interna y máxima diferenciación externa. La clasificación de áreas espaciales mediante técnicas jerárquicas basadas en distancias multivariadas, sigue una rutina como la que se muestra en la [figura 1](#)

Donde el algoritmo considera las distancias multivariantes entre los atributos seleccionados para cada área, ejecuta una iteración para encontrar en cada paso las menores distancias multivariantes (acotadas en la figura bajo el término  $w$ ) hasta optimizar este parámetro, que se devuelve mediante un historial de clasificación y con el la decisión del número de conglomerados de similaridades multivariantes a utilizar y finalmente tras el análisis de la obtención del dendograma se decide las pertenencias de los conjuntos de áreas internamente similares y externamente distantes.

Por último se calcularon estadísticas de indicadores de daños de salud sobre estos estratos. Para lograr una estabilidad adecuada en las cifras de eventos de salud se trabajó con los casos acumulados de 5 años de observaciones, en el caso de las enfermedades transmisibles se utilizó la tasa ponderada de incidencia y en el caso de la morbilidad y mortalidad de enfermedades no transmisibles se utilizó la tasa acumulada y ponderada de prevalencia en la población. Como excepción y dado su carácter no curable, el VIH fue tratado también de esta manera. Estos procedimientos permitieron minimizar los efectos debidos a las diferencias de las poblaciones en riesgo. De esta forma las estadísticas de daño se representaron mediante centroides del estrato, representativos de la proporción de eventos ocurridos en la población contenida en cada estrato. Teniendo en cuenta que los estratos contruidos tienen carácter ordinal con respecto a condiciones contextuales, lo que se pondrá en evidencia al observar la caracterización multivariante de ellos, se utilizaron los valores de significación obtenidos por regresión para evaluar la calidad del ajuste y la característica de la relación estrato-evento.

## RESULTADOS

Los resultados de la clasificación se muestra en la [tabla 1](#) y en la [figura 2](#).

La siguiente descripción da cuenta de las características fundamentales de estos estratos:

Estrato I: Municipios rurales, con baja densidad poblacional, alta relación de dependencia, población relativamente joven, baja concentración urbana y elevado número de asentamientos poblacionales, con baja o inexistente cobertura de servicios de agua potable, con muy baja o inexistente cobertura de disposición de desechos líquidos, base económica agroforestal, industria local rudimentaria, escaso o nulo desarrollo de los servicios y el turismo internacional, condiciones de vivienda mayoritariamente malas, acceso vial limitado y población dispersa. Presencia de áreas verdes con destino deportivo o recreacional, de forma aislada.

Estrato II: Municipios rur-urbanos, con densidad poblacional media, relación de dependencia equilibrada, población envejecida, baja concentración urbana (superior al estrato I) y menor número de asentamientos, base económica agropecuaria, industria local rudimentaria, desarrollo bajo de los servicios y en especial los destinados al turismo internacional, con baja cobertura de servicios de agua potable, presente en comunidades aisladas y no en la cabecera municipal, muy baja o inexistente cobertura de la disposición de desechos líquidos, estado general de la vivienda regular, acceso vial aceptable y población agrupada en pequeños, pero numerosos asentamientos poblacionales. Áreas verdes recreacionales relacionadas al paisaje, no estructuradas.

Estrato III: Municipios fundamentalmente urbanos, con baja relación de dependencia, alta concentración urbana y bajo número de asentamientos poblacionales (generalmente uno), base económica industrial, industrias de alcance provincial o nacional, desarrollo medio de los servicios y el turismo, con cobertura alta de servicios de agua potable, presente fundamentalmente en la cabecera municipal, cobertura media de los sistemas de alcantarillados, principalmente en la cabecera municipal, estado de la vivienda mixto, acceso vial múltiple y mixto en su calidad, y población concentrada en una cabecera municipal que no constituye la totalidad de la misma (pequeñas poblaciones satélites). Presencia de áreas verdes recreacionales estructuradas y distribuidas de acuerdo a la población.

Estrato IV: Municipios totalmente urbanos, baja relación de dependencia, alta concentración urbana que alcanza la totalidad del área municipal, con base económica de servicios y turismo, presencia industrial de alcance provincial o nacional, con alta cobertura de los servicios de agua potable y disposición de desechos, población totalmente concentrada en zonas urbanas, estado de la vivienda mixto y variable, que dada su alta concentración demográfica requiere a su vez ser estratificado, vialidad adecuada e interconectada con viales nacionales, presencia de áreas verdes recreacionales distribuidas en la población.

La superposición de estadísticas de indicadores de daños de salud sobre estos estratos, arrojó interesantes resultados, que se muestra numéricamente en la tabla 2.

Tabla 2. Morbilidad por entidades transmisibles

Morbilidad por entidades transmisibles (tasas x 100 000 excepto VIH, 1 000 000)								
Estratos	EDA	Hep.A	TBP	IRA	Leptospiro sis.	Sífilis	Blenorra gia	VIH/SI DA
1 (n=81)	9413, 3	117,4	7,1	32859,2	9,2	140,9	293,2	5,6
2 (n=59)	5644,	168,2	10,2	52828,8	12,9	80,5	170,9	11,1

	8							
3 (n=13)	8614,5	169,6	13,3	42406,2	7,7	116,4	200,7	19,7
4 (n=16)	9073,8	143,8	14,9	38628,4	6,7	112,6	341,1	42,5
Total	7513,5	141,6	9,4	41111,4	10,1	115,3	247,9	12,1
p	0,692	0,115	0,002	0,492	0,330	0,459	0,068	0,000
Caracterización de la regresión	No	No	Directa	No.	No	No	No	Directa
Morbilidad por entidades no transmisibles seleccionadas (tasas x 100 000 excepto HTA, tasa x 100)								
Estratos	HTA	Asma bronquial	Diabetes mellitus	Cardiovasculares				
1 (n=81)	7,77	55,4	12,0	10,9				
2 (n=59)	9,92	68,1	19,4	20,7				
3 (n=13)	10,3	72,3	20,4	17,2				
4 (n=16)	12,0	87,9	32,0	27,3				
Total	9,2	64,2	17,1	16,4				
p	0,014	0,019	0,038	0,108				
Caracterización de la regresión	Directa	Directa	Directa	No				
Mortalidad por entidades no transmisibles (tasas x 100 000)								
Estratos	Corazón	Cáncer	ECV	Accidentes	Suicidio	Diabetes	Asma bronquial	
1 (n=81)	149,3	118,9	54,0	39,7	18,4	9,6	2,4	
2 (n=59)	197,2	127,8	68,8	43,8	18,9	14,5	2,2	
3 (n=13)	171,6	131,5	71,9	43,1	17,6	15,4	1,3	
4 (n=16)	131,8	150,8	74,3	42,9	12,6	21,2	5,8	
Total	166,1	125,9	62,5	41,7	17,9	12,8	2,6	
p	0,271	0,038	0,003	0,127	0,091	0,030	0,384	
Caracterización de la regresión	No	Directa	Directa	No	Inversa	Directa	No	
Indicadores del programa materno infantil (MI, tasa x 1 000 NV, IBP, %)								
Estratos	MI	IBP						
1 (n=81)	6,8	6,4						
2 (n=59)	7,3	6,5						
3 (n=13)	6,0	6,9						

4 (n=16)	7,9	6,8						
Total	7,0	6,5						
p	0,432							
	0,606							
Caracterización de la regresión	No	No						

EDA: enfermedad diarreica aguda,  
Hep. A.: hepatitis viral tipo A, TBP: tuberculosis pulmonar, IRA: infecciones respiratorias agudas, HTA: hipertensión arterial, ECV: enfermedad cerebrovascular, MI: mortalidad infantil, NV: nacidos vivos, IBP: índice de bajo peso.

## DISCUSION

Los ejes que dieron paso a la clasificación de los municipios permitieron discernir una estructura contextual útil para explicar la distribución y tendencia de la mayoría de los eventos estudiados.

Por orden de estabilidad, la mortalidad por eventos no transmisibles se comportó como el mejor indicador de la estructura espacial contextual, en particular la mortalidad por cáncer, enfermedad cardiovascular (ECV) y diabetes mellitus. En segundo lugar la morbilidad de eventos no transmisibles mostró gran estabilidad y consistencia frente a los estratos contextuales, en particular la hipertensión arterial (HTA), el asma bronquial y la diabetes mellitus.

Los eventos transmisibles mostraron una menor consistencia con la estructura de los estratos, pero esta relación fue particularmente fuerte y consistente para el VIH y la tuberculosis pulmonar (TBP). *Gould* reseña que analizar el modo como se extiende la enfermedad puede revelar cómo está estructurado el espacio.<sup>6</sup>

Los indicadores del Programa Materno Infantil se comportaron de manera independiente a la estructura espacial contextual conformada por los estratos. Este hallazgo es particularmente importante y significativo porque es indicativo de que las acciones del programa son capaces de "eliminar" al espacio y al contexto como elementos significativos, es decir, dejan de ser determinantes en los resultados.

Es en una reflexión como esta donde se puede ver claro que las variables se comportan como "capas" provenientes de diferentes entornos: físico, natural, antrópico, social, pero si las acciones conscientes (programas, intervenciones, indicaciones) que a su vez están mejor o peor definidas según el "estado del arte" asociado de cada evento en particular, llegan en un entorno de equidad a toda la población, muchos eventos evitables pueden reducirse, aplanando las diferencias espaciales y contextuales. La independencia de los indicadores del Programa Materno Infantil de los estratos espacio-contextuales cubanos es un excelente ejemplo del carácter determinista de las acciones del programa sobre los resultados. El espacio ha sido aquí "borrado" mediante una combinación de conocimiento y dominio de variables determinantes y acciones efectivas, equitativas y universales.



La búsqueda de similaridades espaciales basada en el tratamiento multivariante es una herramienta útil que aporta elementos complementarios a los estudios epidemiológicos individualizados. El uso de unidades ambientales no convencionales, o la construcción de unidades espaciales basadas en la integración de áreas con similaridades multivariadas que toman en cuenta variables clave, representa una alternativa metodológica para comprender mejor el proceso de enfermar incorporando el contexto. Este enfoque es flexible, independiente de la unidad espacial básica del análisis, refractario a los problemas de análisis de áreas pequeñas y del cambio de la unidad espacial de análisis (MAUP), así como refractario a los problemas de contigüidades espaciales, ya que crea un macroespacio para fines analíticos ulteriores, así como permite estructurar el espacio más allá de las fronteras, por lo que crea un marco contextual relativamente homogéneo y estable. Si la clasificación multivariante es exitosa, el resultado es la creación de un constructo de riesgo que se comporta como una variable ordinal y se espera una correlación de este constructo con los indicadores de daño sensibles a él, lo que a su vez permite identificar indicadores de la estructura contextual espacial. Estas herramientas son sin dudas fundamentales para la planificación y evaluación de intervenciones en salud pública y en general, para la gestión de políticas de salud.

## REFERENCIAS BIBLIOGRÁFICAS

1. Elliot P, Wakefield J, Best N, Briggs D. Spatial Epidemiology. Methods and Applications. Oxford: University Press; 2001.
2. Haining R. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge: Cambridge University Press; 2000.
3. Apitz de Parra MJ, Escalona FF, Taborda R. Cluster analysis in the determination of homogeneous regions based on the production of maize (corn) and sorghum in the zulia region. Rev Fac Agron (LUZ). 1998;15:608-20.
4. Bara T. Gis-Based regionalization of natural landscape using derived landcover occurrence probabilities. In: GIS/LIS, editors. Proceedings of the 1994 anual conference and exposition. Arizona, EE.UU.: GIS/LIS; 1994.p.34-43.
5. Alegret M, Spiegel J, Yassi A. Importance of Stratification and Territorial Ordering in Prospection. Prediction and Surveillance in Public Health. Presented at 4th Annual Meeting of the CCGHR. Ottawa: CCGHR; 2004.
6. Gould P. Pensamientos sobre la Geografía. Geocrítica.1986; Año XII(68).

Recibido: 24 de Julio de 2007.

Aprobado: 11de octubre de 2007.

Milagros Alegret Rodríguez. Unidad de Análisis y Tendencias en Salud. Centro Provincial de Higiene y Epidemiología. Villa Clara, Cuba. E-mail: [malegret@dps.vcl.sld.cu](mailto:malegret@dps.vcl.sld.cu)

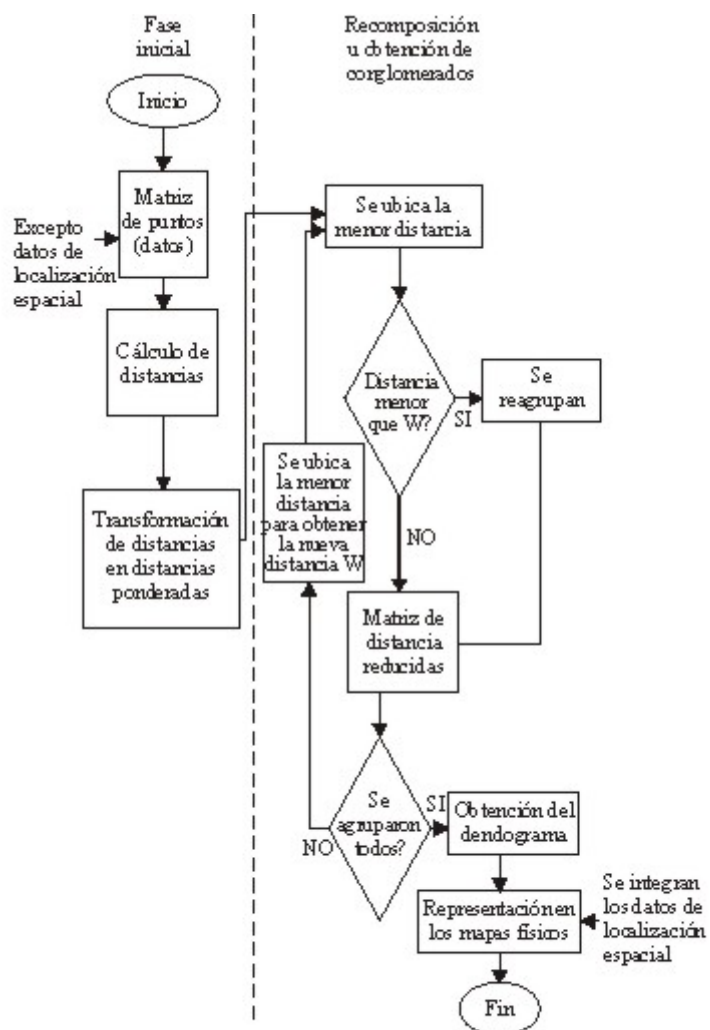


Fig.1. Diagrama de flujo de los procesos de conformación de conglomerados espaciales.

Tabla 1. Distribución de las municipalidades cubanas en los estratos

Estrato	No. de municipios	%
I	81	47,9
II	59	34,9
III	13	7,7
IV	16	9,5
Total	169	100

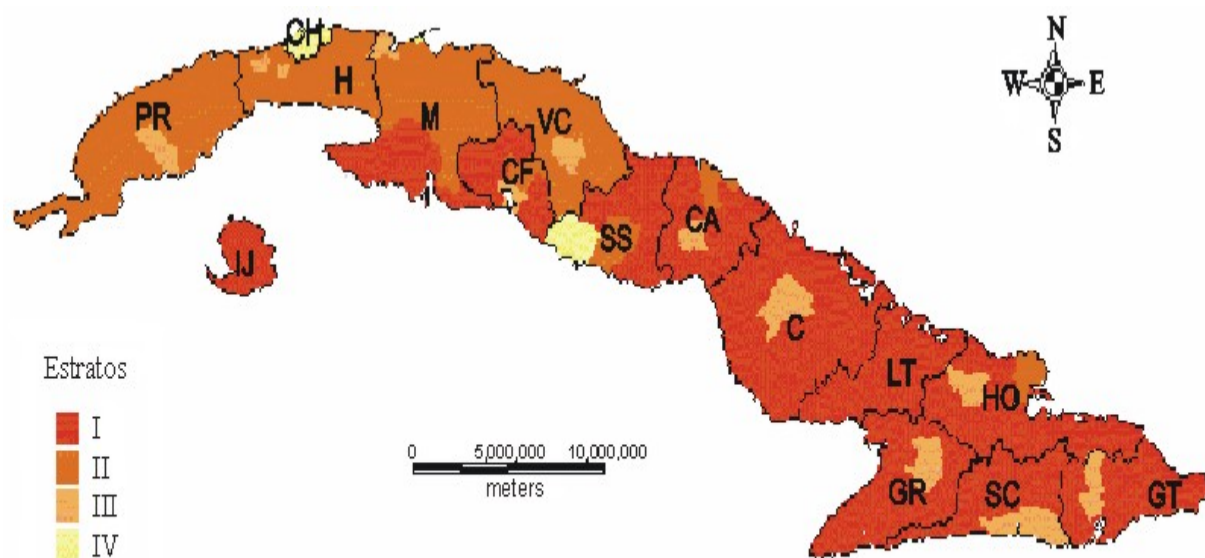


Fig. 2. Ubicación geográfica de los estratos de municipalidades cubanas.