

Censo Económico de Colombia

Documento metodológico

de identificación y delimitación
de áreas de concentración
económica rural (ACER)

2021



El futuro
es de todos

Gobierno
de Colombia

DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA -DANE-

***DIRECCIÓN DE METODOLOGÍA Y
PRODUCCIÓN ESTADÍSTICA -DIMPE-***

CENSO ECONÓMICO DE COLOMBIA

**DOCUMENTO METODOLÓGICO DE
IDENTIFICACIÓN Y DELIMITACIÓN
DE ÁREAS DE CONCENTRACIÓN
ECONÓMICA RURAL (ACER)**

2021

**DEPARTAMENTO ADMINISTRATIVO
NACIONAL DE ESTADÍSTICA
(DANE)**

JUAN DANIEL OVIEDO ARANGO

Director

RICARDO VALENCIA RAMIREZ

Subdirector

MARÍA FERNANDA DE LA OSSA ARCHILA

Secretaria General

DIRECTORES TÉCNICOS

HORACIO CORAL DIAZ

Dirección de Metodología y Producción Estadística

ANGELA VEGA LANDAETA

Dirección de Censos y Demografía

JULIETH ALEJANDRA SOLANO VILLA

Dirección de Regulación, Planeación,
Estandarización y Normalización

JUAN PABLO CARDOSO TORRES

Dirección de Síntesis y Cuentas Nacionales

SANDRA LILIANA MORENO MAYORGA

Dirección de Geoestadística

MAURICIO ORTIZ GONZÁLEZ

Dirección de Difusión, Mercadeo y
Cultura Estadística

© DANE, 2021

Prohibida la reproducción total o parcial sin permiso o auto-rización del Departamento Administrativo Nacional de Estadística, Colombia.

TABLA DE CONTENIDO

INTRODUCCIÓN	9
1. MARCO TEÓRICO	11
1.1. DEFINICIÓN DE ÁREA DE CONCENTRACIÓN ECONÓMICA (ACER)	11
2. METODOLOGÍA	13
2.1. DESCRIPCIÓN GENERAL	13
2.2. ALCANCE DE LA METODOLOGÍA	15
2.3. DESARROLLO METODOLÓGICO	15
2.3.1 Identificación y selección de insumos	15
2.3.2 Preprocesamiento de fuentes de información	20
2.3.3 Selección de puntos por municipio y reproyección de capas a coordenadas planas	26
2.3.4 Cálculo de la mediana de las distancias de los establecimientos a su vecino más cercano	27
2.3.5 Geoprocesos de generación de polígonos de referencia de las ACER	29
2.3.6 Validación de ACER mediante variables exógenas	33
2.3.7 Validación y codificación de ACER	54
2.3.8 Indicadores de importancia relativa de las ACER	56
3. RESULTADOS	61
3.1. RESULTADOS DEL MUNICIPIO DE UBATÉ (CUNDINAMARCA)	61
4. CONCLUSIONES	68

LISTA DE TABLAS

Tabla 1 Fuentes de información oficiales en la metodología de identificación y delimitación de ACER	17
Tabla 2 Fuentes de información alternativas en la metodología de identificación y delimitación de ACER	18
Tabla 3 Fuentes de información eventuales en la metodología de identificación y delimitación de ACER	19
Tabla 4 Códigos de uso de edificación de interés económico en la capa CONSTRUCCION_P	22
Tabla 5 Total de registros georreferenciados (puntos) por fuente de información	25
Tabla 6 Total de registros georreferenciados (puntos) por fuente de información	36
Tabla 7 Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3,4,5	38
Tabla 8 Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3,4	38
Tabla 9 Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3	39
Tabla 10 Distribución de distancias de establecimientos económicos al cuerpo de agua más cercano según región	39
Tabla 11 Rangos de valores de pendiente	41
Tabla 12 Vectores de pesos por escenario	47
Tabla 13 Resultados de umbrales con el método de intervalos iguales para diferentes municipios	51
Tabla 14 Parámetros del municipio de Ubaté (Cundinamarca)	61
Tabla 15 Algunos resultados de validación de ACER en el municipio de Ubaté (Cundinamarca)	66
Tabla 16 Resultados de indicadores de importancia de ACER para el municipio de Ubaté (Cundinamarca)	67

LISTA DE FIGURAS

Figura 1 Diagrama de flujo de la metodología de identificación y delimitación de las ACER	14
Figura 2 Distribución geográfica de la capa consolidada de fuentes de información GEO_FUENTES_CONSOLIDADAS_2020	26
Figura 3 Diagrama de caja y bigotes de las distancias de los establecimientos económicos a su vecino más cercano para el municipio de Ubaté (Cundinamarca)	28
Figura 4 Grilla hexagonal para el municipio de Ubaté (Cundinamarca) y establecimientos económicos, donde la altura del hexágono viene dada por $Me = 207,7659$ m	30
Figura 5 Áreas de influencia (Buffer) de los establecimientos económicos del municipio de Ubaté (Cundinamarca). El radio del buffer corresponde a $Me = 207,7659$ m	31
Figura 6 Superposición de polígonos LISA y Buffer en el municipio de Ubaté (Cundinamarca) obtenidos sobre los establecimientos económicos	32
Figura 7 Esquema general de la evaluación multicriterio.	34
Figura 8 Tipos de vía definidos en la cartografía base del IGAC, escala 1:100.000	37
Figura 9 Reclasificación de luces nocturnas del municipio de Fusagasugá por medio de intervalo geométrico	40
Figura 10 Ejemplo de matriz de comparación por pares (MCP)	42
Figura 11 Escala fundamental de comparación por pares	42
Figura 12 Suma por columna de la MCP	43
Figura 13 . MCP normalizada	44
Figura 14 Vector de pesos con las ponderaciones o porcentajes	44
Figura 15 Producto matricial y cociente de vectores	45
Figura 16 Media geométrica de MCP y vector de pesos para escenario 1	46
Figura 17 Media geométrica de MCP y vector de pesos para escenario 2	46
Figura 18 Media geométrica de MCP y vector de pesos para escenario 3	47
Figura 19 MCP de clases dentro de las variables y vector de pesos	48
Figura 20 Vector de puntajes dentro de cada variable	48
Figura 21 Rangos de las variables y puntajes en el municipio de Fusagasugá	49

Figura 22 Ilustración de la suma ponderada de imágenes ráster en el municipio de Fusagasugá	50
Figura 23 Resultado del cálculo de la suma ponderada de las variables en el municipio de Fusagasugá	50
Figura 24 Resultado de la reclasificación de la suma ponderada en el municipio de Fusagasugá	53
Figura 25 . Intersección de ACER con área óptima calculada mediante AHP	53
Figura 26 Ejemplo de código para área de concentración económica rural (ACER)	55
Figura 27 Ubicación de establecimientos económicos en clase 3 (rural disperso) en el municipio de Ubaté (Cundinamarca)	55
Figura 28 Ubicación de áreas de influencia o buffer en el municipio de Ubaté (Cundinamarca)	63
Figura 29 Ubicación de clusters y outliers de establecimientos económicos en Ubaté (Cundinamarca)	63
Figura 30 Ubicación de polígonos de referencia de ACER en Ubaté (Cundinamarca)	64
Figura 31 Imágenes ráster para calcular el área óptima o potencial de actividad económica rural en Ubaté.	64
Figura 32 Imágenes ráster de resultado de la ponderación y selección del área óptima de actividad económica en el rural disperso	65
Figura 33 Área óptima intersectada con ACER del municipio de Ubaté	65
Figura 34 Resultado de estructuración de las ACER del municipio de Ubaté (Cundinamarca)	67

LISTADO DE SIGLAS

ACER	Áreas de Concentración Económica Rural
CNPV 2018	Censo Nacional de Población y Vivienda 2018
DANE	Departamento Administrativo Nacional de Estadística
DNP	Departamento Nacional de Planeación
IGAC	Instituto Geográfico Agustín Codazzi
INVIAS	Instituto Nacional de Vías
INVIMA	Instituto Nacional de Vigilancia de Medicamentos y Alimentos
MEN	Ministerio de Educación Nacional
MGN	Marco Geoestadístico Nacional
MINCIT	Ministerio de Comercio, Industria y Turismo
MINSALUD	Ministerio de Salud
OSM	Open Street Map

INTRODUCCIÓN

El Departamento Administrativo Nacional de Estadística (DANE) en cumplimiento de lo establecido en el documento CONPES 3956 de 2019 adelanta el proceso de diseño del Censo Económico de Colombia (CE) que se realizará en el año 2022. En este proceso se estructura técnicamente y se detallan el conjunto de métodos de los cuales se hará uso para el desarrollo de esta operación estadística.

Los censos económicos, al ser las operaciones estadísticas más usuales de recopilación de datos para describir las características y actividades de las unidades económicas, constituyen la fuente de datos más completa acerca de la estructura de la economía permitiendo, entre otros aspectos, recolectar información actualizada sobre el número, la distribución espacial y las características económicas y temáticas relacionadas de los establecimientos de industria, comercio, servicios, construcción y transporte, y contar con insumos para el cálculo de las variables macroeconómicas utilizadas en las cuentas nacionales y regionales del país¹.

Desde 1990, año del último Censo Económico Nacional y Multisectorial, no se realiza un ejercicio semejante a nivel nacional. No obstante, el Censo de Población del 2005 incluyó un módulo relacionado con unidades económicas, mientras que el Censo Nacional Agropecuario de 2014 diferenció las actividades económicas del sector. El Censo de Población y Vivienda del 2018 realizó un conteo previo y una verificación de establecimientos económicos diferenciados por sector. Así mismo, se han realizado algunos ejercicios a nivel territorial como el Censo Económico de Cundinamarca de 1999, el Censo Económico de Santiago de Cali (2005) o la Encuesta a Establecimientos Económicos de Bogotá de 2017².

Dadas las implicaciones de realizar un Censo Económico después de más tres décadas (siguiendo la recomendación del documento CONPES 3956 sobre política de Formalización Empresarial)³ resulta necesario diseñar y emprender acciones que permitan aproximarse de la manera más precisa y representativa a

¹ Departamento Administrativo Nacional de Estadística (2019). Censo Económico de Colombia, documento metodológico preliminar resultante consulta pública. Recuperado de <https://www.dane.gov.co/files/investigaciones/censo-economico/documento-conceptual/documentos/FASE-III-documento-metodologico-preliminar-censo-economico.pdf>

² Ibid.

³ Departamento Nacional de Planeación (2019). Documento CONPES 3965 Política de formalización empresarial. Recuperado de <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3956.pdf>

la dimensión, estructura y distribución de la actividad económica en Colombia; con el objetivo de optimizar la asignación de recursos y la realización del operativo censal. Al respecto, se tiene proyectado un alcance geográfico del 100% para cabeceras municipales y centros poblados (clases 1 y 2) y un alcance parcial en el rural disperso (clase 3).

Para la clase 3 se ha definido el concepto de áreas de concentración económica rural (en adelante ACER) con el fin de realizar el operativo del Censo Económico tanto en dichas áreas y en las unidades económicas de inclusión forzosa; dada la complejidad de una operación de barrido total por los altos costos que conlleva y por la necesidad de focalizar el esfuerzo del operativo censal en sitios y establecimientos representativos de la actividad económica en esta zona. Por otra parte, en la clase 3 no se contempla realizar operaciones de conteo de establecimientos económicos.

Este documento presenta la metodología planteada, en el Grupo Interno de Trabajo de Investigación y Desarrollo de la Dirección de Geoestadística (DIG) del DANE, para la identificación y delimitación de las ACER, así como el proceso de validación con variables exógenas y de codificación para su inclusión en el marco censal⁴.

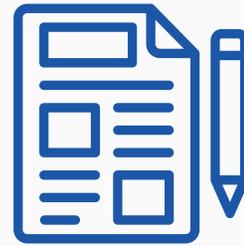
El proceso que aquí se expone se enmarca en la definición de Diseño Estadístico del Documento Metodológico del Censo Económico: "La definición, la construcción y/o la formulación del Marco Estadístico, el universo, la población, las unidades estadísticas, los periodos y el diseño muestral (...) Para el caso de operaciones estadísticas por censo, se diseña el marco censal"

En primer lugar, se presenta la definición de ACER fundamentada en la literatura sobre geografía económica. A continuación, se encuentra el proceso metodológico de identificación y delimitación de las ACER, que incluye la selección de insumos o fuentes de información, el preprocesamiento de las fuentes, el flujo de procesos geoespaciales para la generación de los polígonos de referencia de las ACER con la respectiva explicación de los algoritmos empleados, la validación y priorización de ACER empleando variables exógenas y análisis multicriterio, la delimitación y codificación dentro del Marco Geoestadístico Nacional (MGN)⁵ y los indicadores propuestos para establecer la importancia de las ACER a nivel municipal. Finalmente, se presenta un ejemplo de identificación, delimitación y codificación de ACER para el municipio de Ubaté (Cundinamarca), con sus respectivos indicadores.

⁴ Sistema Estadístico Nacional - Colombia (2020). Guía para la elaboración del documento metodológico de operaciones estadísticas. Recuperado de https://www.sen.gov.co/files/sen/lineamientos/Guia_Documento_Metodologico.pdf

⁵ Es el sistema para referenciar la información estadística a su localización geográfica. Está constituido por áreas geoestadísticas (departamentos, municipios, cabeceras municipales, centros poblados, rural disperso, entre otras) delimitadas principalmente por accidentes naturales y culturales, identificables en terreno y su respectiva codificación

1. MARCO TEÓRICO



1.1 Definición de área de concentración económica (ACER)

La Nueva Geografía Económica, definida por Krugman y Fujita⁶, trata la disposición de las actividades económicas en espacios geográficos a distintos niveles, estudiando las fuerzas que conllevan tanto la concentración como la dispersión o separación de dichas actividades⁷. Según Krugman, factores espaciales como la vecindad, la cercanía y la menor distancia entre las personas y la actividad económica, generan ventajas intrínsecas que favorecen la productividad y la competitividad. Lo anterior conforma un modelo que explica las dinámicas de agrupación de establecimientos, considerando las fuerzas centrípetas o elementos que por sus ventajas implican una concentración de actividades económicas; y las fuerzas centrifugas, las cuales conllevan una dispersión de las actividades y generalmente tienen un origen externo⁸.

Entre las ventajas de las agrupaciones o aglomeraciones de establecimientos económicos se encuentran: la generación de empleo para mano de obra especializada, la producción y divulgación de información referente a las actividades económicas realizadas, el impacto positivo en las comunidades aledañas o la reducción de costos de transporte⁹.

Para entender la dinámica de la actividad económica en el contexto geográfico se han desarrollado estudios tanto a nivel nacional como internacional. Propuestas en Colombia como la detección de aglomeraciones económicas en Bogotá empleando herramientas de estadística espacial¹⁰, o trabajos a nivel internacional como el análisis de la agrupación espacial de la economía rusa¹¹, los estudios de patrones de

⁶ Fujita, M., & Krugman, P. (2004). La nueva geografía económica. *Papers in regional science*, 83(1), 177-206.

⁷ DANE (2020). Documento de conceptualización de área de concentración económica.

⁸ Ibid.

⁹ Ibid.

¹⁰ Gaitan, J. (2013). Incidencia de las economías de aglomeración en los sectores reales localizados en el área urbana de Bogotá. *Ensayos sobre política económica*, 31(70).

¹¹ Aleksandrova, E., Behrens, K., & Kuznetsova, M. (2019). Manufacturing agglomeration in a transition country: Evidence from Russia. *Journal of regional science*, 60, 88– 128. doi:10.1111/jors.12436

agrupamiento de diferentes ramas económicas en regiones rurales de Suecia¹², o la aplicación de metodologías de análisis estadístico espacial a la localización de empresas dirigidas por mujeres en Grecia¹³; son experiencias que muestran cómo la distribución espacial de establecimientos económicos no es de ningún modo aleatoria o circunstancial, sino que tiende a formar aglomeraciones susceptibles de ser analizadas a través de métodos estadísticos, de acuerdo con la manera de considerar el espacio de los eventos. Si es discreto, se emplean métodos tales como el coeficiente Gini de localización, el área de influencia (buffer), el escaneo estadístico y los índices locales de autocorrelación espacial (LISA); para un espacio continuo se utiliza la función K de Ripley o las estructuras de redes subyacentes¹⁴.

Los trabajos anteriores también evidencian que la distribución espacial está influida por factores como la facilidad de acceso a vías de transporte y la disponibilidad de recursos físicos y humanos¹⁵.

Cuando la agrupación espacial de establecimientos económicos se presenta en la zona rural la superficie donde se concentran se denomina área de concentración económica rural (o ACER), que se caracteriza por procesos de especialización de sectores económicos específicos, o de diversificación por la presencia de establecimientos pertenecientes a diferentes renglones.

Se define formalmente el área de concentración económica rural (ACER) como un área delimitada en el rural disperso (clase 3)¹⁶, que contiene establecimientos de uno o diferentes sectores económicos, los cuales están aglomerados, o concentrados espacialmente, con un número mínimo de establecimientos.

La concentración espacial se determina a partir de la utilización de técnicas estadísticas y teniendo en cuenta condiciones inherentes a las fuentes disponibles tales como distancia, vecindad o cercanía¹⁷.

¹² Vânia Ceccato Lars, O. P. (2002). Dynamics of rural areas: an assessment of clusters of employment in Sweden. *Journal of Rural Studies*, 18, 49-63. doi:[https://doi.org/10.1016/S0743-0167\(01\)00028-6](https://doi.org/10.1016/S0743-0167(01)00028-6)

¹³ Bersimis, S., Chalkias, C., & Anthopoulou, T. (2012). Detecting and interpreting clusters of economic activity in rural areas using scan statistic and LISA under a unified framework. *Applied Stochastic Models in Business and Industry*, 573-587. doi:10.1002/asmb.2003

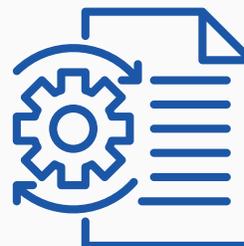
¹⁴ Departamento Administrativo Nacional de Estadística (2020), Norma Técnica de la Calidad del Proceso Estadístico. Requisitos de Calidad para la Generación de Estadísticas. Recuperado de: <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

¹⁵ Ibid.

¹⁶ Delimitación geográfica definida por el DANE para fines estadísticos, comprendida entre el perímetro censal de las cabeceras municipales y de los centros poblados, y el límite municipal. Se caracteriza por la disposición dispersa de viviendas y de explotaciones agropecuarias existentes en ella. Corresponde al territorio que no forma parte ni de la cabecera municipal (clase 1) ni de los centros poblados (clase 2). Se caracteriza por objetos y elementos relacionados con la agricultura, predios de descanso o recreo, usos mineros o extractivos. El número de unidades residenciales por área es menor a las zonas urbanas (Departamento Administrativo Nacional de Estadística, DANE, 2018).

¹⁷ Departamento Administrativo Nacional de Estadística (2020), Construcción interna propia Mesas DIMPE.

2. METODOLOGÍA



2.1. Descripción general

La metodología de identificación y delimitación de las ACER consta de los siguientes macroprocesos:

Identificación de los insumos: Consiste en la consulta y selección de las fuentes de información que proveen la ubicación, y otras características de los establecimientos económicos, teniendo en cuenta cinco criterios: pertinencia, disponibilidad, vigencia, calidad y coherencia. Estos insumos son esenciales para el modelamiento geoespacial y la validación posterior.

Preprocesamiento de fuentes de información: Se realiza una revisión de cada una de las fuentes de información, lo cual implica operaciones de estandarización, limpieza y comparación para eliminar posibles duplicados, para luego proceder a la integración (fusión) con las demás fuentes y generar una sola capa geográfica consolidada.

Selección de puntos por municipio y reproyección de capas a coordenadas planas: Una vez las fuentes se consolidan en una sola capa geográfica, en donde es posible identificar la fuente original mediante un atributo, se realiza un geoproceso que consiste en seleccionar espacialmente los establecimientos económicos por municipio, verificar que cada registro corresponda efectivamente al municipio donde se encuentra, y reproyectar la capa resultante

en las respectivas coordenadas planas MAGNA SIRGAS¹⁸ del municipio, ya que así lo requieren las etapas posteriores. También se reproyecta el polígono del límite del municipio, proveniente del repositorio del Marco Geoestadístico Nacional (MGN) en su última actualización.

Geoprocesos de generación de polígonos de referencia de las ACER: Una vez obtenidas las capas de puntos y el límite municipal en sus coordenadas planas correspondientes, se ingresan en el geoproceso implementado en el software ArcMap® 10.5.1 para la identificación y delimitación de las ACER a través de dos algoritmos: Área de Influencia (Buffer) e Índices Locales de Autocorrelación Espacial (LISA). Cada uno de los algoritmos tiene unos parámetros específicos que se tratarán en detalle en este documento. Como resultado de esta etapa se obtienen los polígonos de referencia de las ACER por municipio.

Validación de las ACER a través de variables exógenas jerarquizadas con análisis multicriterio: Los polígonos de referencia de las ACER se intersectan con una nueva capa denominada Área Óptima o Área Potencial que

¹⁸ Marco Geocéntrico Nacional de Referencia. Sistema de Referencia Geocéntrico para las Américas

corresponde a aquella zona del municipio con alta importancia económica dentro del rural disperso. Esta área está determinada mediante la suma ponderada de cinco (5) variables exógenas a la ubicación de los establecimientos económicos, donde la participación o porcentaje de cada variable se calcula mediante un análisis multicriterio o jerárquico (AHP). Finalmente se seleccionan los polígonos de referencia que se crucen con el área potencial, junto con algunos de los que quedan por fuera de dicha área, de acuerdo a unos criterios definidos de inclusión.

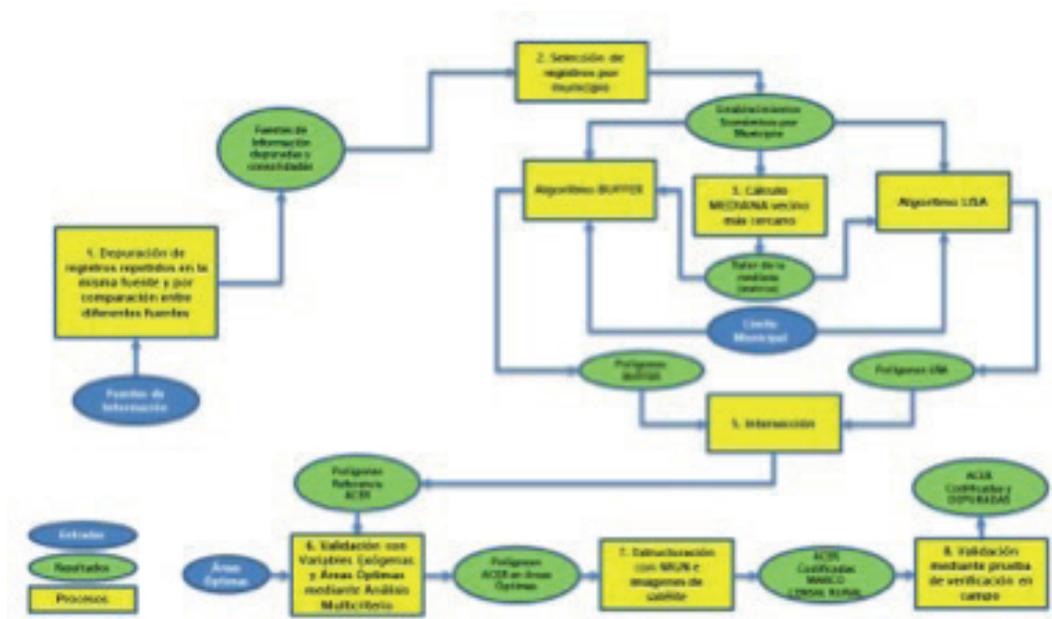
Delimitación con respecto al MGN y codificación de ACER: Los polígonos de referencia de las ACER deben ajustarse a los niveles geográficos del Marco Geoestadístico Nacional (MGN) con el fin de otorgar a cada área un código único. Durante este proceso debe evaluarse si las áreas calculadas en realidad presentan concentración de establecimientos económicos mediante comparación con imágenes de satélite y teniendo como referencia las áreas óptimas calculadas a través del análisis multicriterio. Posteriormente, se realiza la delimitación de las ACER teniendo en cuenta límites naturales, límites prediales,

polígonos del Marco Maestro Rural Agropecuario (MMRA) y límites de secciones rurales.

Evaluación de las ACER: por parte de las dependencias interesadas del DANE. Las ACER delimitadas y codificadas serán validadas en campo a través de una muestra representativa de municipios a nivel departamental. Los resultados serán evaluados junto con la Dirección de Metodología y Producción Estadística (DIMPE) y el Área de Logística, y se tomarán las medidas de ajuste pertinentes para su estructuración final e inclusión en el marco censal, con el fin de ser insumo para el operativo del Censo Económico. Para ello, la DIG elaborará una propuesta de diseño de la prueba, coordinará su desarrollo, recopilará los resultados y ajustará la metodología según las observaciones.

El diagrama de flujo de la metodología de identificación y delimitación de ACER con la secuencia de los macroprocesos ejecutados (amarillo), las entradas y los productos obtenidos hasta tener las ACER estructuradas (verde), se puede observar en la Figura 1

Figura 1. Diagrama de flujo de la metodología de identificación y delimitación de las ACER



Fuente: DANE. Censo Económico

2.2. Alcance de la metodología

La metodología de identificación y delimitación de ACER tiene un alcance preciso, definido en los siguientes puntos:

- La identificación y delimitación de ACER por medio de los algoritmos de Buffer y de LISA, así como el proceso de estructuración y codificación, considera únicamente la localización de los establecimientos económicos. No tiene en cuenta variables que puedan caracterizar la importancia económica de los establecimientos (número de empleados, ingresos, etc.), debido a que estos atributos no se tienen para la mayoría de las fuentes de información.
- Debido a que cada municipio presenta unas características socioeconómicas distintas, el ejercicio se implementa separadamente para cada uno. No obstante, la metodología contempla un análisis en los límites municipales.
- Con el fin de priorizar las ACER obtenidas por municipio, se aplica un procedimiento de validación y eliminación de ACER empleando variables exógenas. No obstante, este proceso tampoco es definitivo por su naturaleza determinística, lo que implica la revisión de las ACER restantes.
- Las fuentes de información tienen implícita una incertidumbre de ubicación, completitud y coherencia que inciden en el resultado final.
- El procesamiento de fuentes se realiza a nivel de localizaciones puntuales pero la delimitación se ajusta al Marco Geoestadístico Nacional para definir el marco operativo censal rural.
- La Dirección Técnica de Geoestadística entregará las ACER por municipio, sobre las cuales las áreas de logística y operativo definirán aquellas en las que se hará la recolección de información.

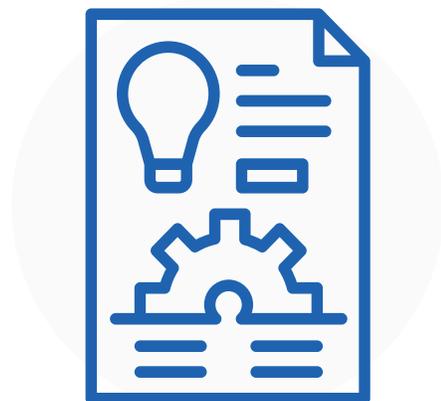
2.3. Desarrollo metodológico

2.3.1. Identificación y selección de insumos

2.3.1.1 Criterios de selección de insumos

Para la selección de los insumos se tienen en cuenta los criterios de pertinencia, disponibilidad, vigencia y coherencia de la información, los cuales se detallan a continuación:

Pertinencia: La fuente de información debe proporcionar registros pertinentes para la identificación y delimitación de ACER. Es decir, dichos registros deben estar localizados en el rural disperso (clase 3).



Disponibilidad: La fuente de información debe estar disponible para su consulta y utilización en todo momento, incluyendo sus actualizaciones posteriores.

Vigencia: Es importante tener en cuenta la vigencia, entendida como la fecha en la que se genera la información, o su fecha más reciente de actualización. Se descartan fuentes con vigencias o actualizaciones mayores a cinco (5) años.

Coherencia: Se refiere al grado en que están lógicamente conectados los conceptos utilizados, las metodologías aplicadas y los resultados producidos por la operación.



2.3.1.2 Fuentes oficiales de información seleccionadas

Se entienden como fuentes oficiales de información las entidades públicas o privadas de carácter nacional, regional o local, cuyos datos están abiertos al público, o bien requieren de un proceso interno de solicitud.

Las fuentes oficiales, con su respectiva información sobre vigencia, descripción, tipo de geometría y justificación, se encuentran en la Tabla 1.

2.3.1.3 Fuentes alternativas de información seleccionadas

Junto a las fuentes oficiales de información, se encuentran fuentes alternativas con información complementaria, proveniente de Internet y capturada a través de algoritmos de Web Scraping o mediante descarga directa.

Las fuentes alternativas con su respectiva información sobre vigencia, descripción, tipo de geometría y justificación, se encuentran en la Tabla 2.

2.3.1.4 Fuentes eventuales de información

Además de las fuentes oficiales y alternativas, existen fuentes provenientes de la base de investigaciones del DANE y de otras entidades, referidas principalmente a operaciones estadísticas sectoriales que se realizan en periodos intercensales. Algunas de estas operaciones constan de formularios georreferenciados.

No se emplean directamente en la metodología por cuanto las operaciones aportan información solo para una parte del territorio. Por esta razón se denominan eventuales y se podrían llegar a utilizar si se requiere información adicional para un municipio o zona determinada.

Las fuentes eventuales con su respectiva información sobre el nombre de la operación estadística y la fuente, se muestran en la Tabla 3.

Tabla 1: Fuentes de información oficiales en la metodología de identificación y delimitación de ACER

Fuente	Base	Vigencia	Descripción	Tipo geometría	Justificación
DANE	CNPV Censo Nacional de Población y Vivienda	2018	Formularios Georreferenciados correspondientes a Clase 3 (Rural Disperso) Usos mixtos Otros Usos	Punto	Es la fuente de información de origen DANE más actualizada. Los formularios georreferenciados permiten diferenciar usos mixtos y usos no residenciales, para determinar ubicación y concentración de establecimientos económicos.
DANE	DEST- Directorio Estadístico Nacional de Empresas	Según disponibilidad	Localización de empresas registradas en el Directorio Estadístico Nacional.	Punto	El DEST es el compendio de varias fuentes oficiales y alternativas (CONFECÁMARAS, einforma.com, etc.), en el figuran las empresas legalmente constituidas en el país.
IGAC	Cartografía Base 1:100.000	2017	Localización puntual de construcciones de interés económico (excluyendo establecimientos educativos).	Punto	Las construcciones tienen una codificación que permite identificar varias de ellas como de interés o importancia económica y complementan la información de las fuentes DANE.
IGAC	Mapas de Ruta Colombia	2018	Localización de establecimientos de interés económico (sitios de interés, hoteles, estaciones de servicio).	Punto	La publicación "Mapas de Ruta Colombia" da información sobre los sitios de interés en 17 rutas que comunican a las principales ciudades del país.
INVIMA	Plantas de Beneficio Animal y Vegetal	2019	Localización de las plantas de beneficio animal y vegetal registradas ante INVIMA.	Punto	Fuente de información sobre actividades económicas del área rural proporcionada por el INVIMA.
MINCIT	Registro Nacional de Turismo RNT	2019	Localización de los establecimientos hoteleros a nivel nacional	Punto	Información complementaria al CNPV 2018, fuente de información sobre actividades económicas en el sector hotelero.
INVIAS	Peajes	2020	Localización de peajes.	Punto	Los peajes, además de ser establecimientos de inclusión forzosa en el Censo Económico, también son generadores de actividad económica en sus cercanías.
MINSALUD	Instituciones Prestadoras de Salud (IPS)	2019	Localización de instituciones prestadoras de salud (IPS).	Punto	El DANE solicitó información actualizada sobre instituciones prestadoras de salud (IPS) en Colombia durante el año 2019. Se entregó una tabla de IPS con datos de cada establecimiento, incluida la georreferenciación. Sector de interés para el Censo Económico.

MEN	Establecimientos Educativos	2019	Localización de establecimientos educativos.	Punto	El DANE solicitó información actualizada sobre instituciones educativas en Colombia durante el año 2019. Se entregó una tabla con datos de cada establecimiento, incluida la georreferenciación. Sector de interés para el Censo Económico.
-----	-----------------------------	------	--	-------	--

Fuente: DANE. Censo Económico

Tabla 2: Fuentes de información alternativas en la metodología de identificación y delimitación de ACER

Fuente	Base	Vigencia	Descripción	Tipo geometría	Justificación
EINFORMA	Base de Datos einforma.co	2019	Localización de la empresa	Punto	La base de datos de einforma contiene 258361 registros, todos ellos georreferenciados, y con atributos importantes como el NIT, lo que permite vincularlo con el DEST.
OPEN STREET MAP	Open Street Map	2020	Localización del establecimiento económico	Punto	Open Street Map es un proyecto de cartografía colaborativa, generada mediante GPS y edición en línea. Su descarga es gratuita. Permitiría encontrar algunos establecimientos económicos no registrados en fuentes oficiales.
PÁGINAS AMARILLAS	Base de Datos Páginas Amarillas	2019	Localización de la empresa	Punto	Páginas Amarillas es una conocida base de datos en papel y en Internet, que aporta aproximadamente 170.000 registros en Colombia, todos geolocalizados.
BOOKING.COM	Base de datos Booking.com	2020	Localización de establecimientos hoteleros	Punto	Booking.com es una plataforma que ofrece información de establecimientos hoteleros a nivel nacional e internacional. En Colombia se encuentran 14.000 registros, extraíbles mediante técnicas de Web Scraping.

Fuente: DANE. Censo Económico

Tabla 3: Fuentes de información eventuales en la metodología de identificación y delimitación de ACER

Operación estadística	Entidad responsable
Estadísticas de los Proyectos Integrales de Desarrollo Agropecuario y Rural – PIDAR.	Agencia de Desarrollo Rural
Estadísticas de los Proyectos y Distritos de Adecuación de Tierras.	Agencia de Desarrollo Rural
Estadística de Monitoreo y Seguimiento RUA Manufacturero en Colombia.	IDEAM
Producción de huevo y pollo.	Federación Nacional de Avicultores - FENAVI
Estadísticas del registro nacional de palmicultores.	Federación Nacional de Cultivadores de Palma de Aceite - FEDEPALMA
Censo de especies animales de importancia económica en Colombia.	Instituto Colombiano Agropecuario - ICA
Información estadística hotelera de los afiliados a COTELCO.	COTELCO
Encuesta anual de Comercio (EAC).	DANE
Censo Nacional Arrocero.	Federación Nacional De Arroceros - Fedearroz
Encuesta Anual Manufacturera (EAM).	DANE
Encuesta anual de servicios (EAS).	DANE
Encuesta ambiental de hoteles (EAH).	DANE
Censo nacional de palma de aceite.	Federación Nacional de Cultivadores de Palma de Aceite Fedepalma
Encuesta ambiental industrial (EAI).	DANE
Componente de insumos y factores asociados a la producción agropecuaria – SIPSA.	DANE
Encuesta nacional agropecuaria (ENA).	DANE
Encuesta mensual de servicios – EMS.	DANE
Cuenta satélite de turismo (CST)	DANE

Cuenta satélite de la agroindustria.	Cuenta Satélite Piloto de la Agroindustria -CSPA-
Encuesta nacional logística (ENL).	Departamento Nacional de Planeación (DNP)
Encuesta nacional de CTI agropecuaria.	Ministerio de Agricultura

Fuente: DANE. Censo Económico

2.3.2. Preprocesamiento de fuentes de información

El resultado final de la aplicación de la metodología de identificación y delimitación de ACER depende de las fuentes de información provistas, dado que se basa totalmente en información secundaria. De este modo, es fundamental la revisión de cada uno de los conjuntos de datos proveniente de las fuentes de información, enfatizando en los siguientes aspectos:

- Los registros en todos los conjuntos de datos deben filtrarse de tal forma que se localicen en el rural disperso (clase 3). Esto puede hacerse mediante una selección por atributos o por localización en un software de análisis de datos espaciales.
- Todas las fuentes de información deben encontrarse inicialmente en coordenadas geográficas MAGNA, con el fin de reproyectar posteriormente los registros seleccionados por municipio en las respectivas coordenadas planas.
- Mediante un análisis textual y comparación de coordenadas, se excluyen los registros de las fuentes que presenten evidencias de duplicación.

Análisis textual

El análisis textual consiste en encontrar registros en varias fuentes que tengan un nombre y actividad económica similar, para luego comparar su ubicación geográfica. De esta manera se garantiza

que no haya repetición del mismo establecimiento en diversas fuentes de información. El análisis textual se realiza empleando rutinas escritas en lenguajes de programación como R o Python.

Comparación de coordenadas

Las coordenadas geográficas de los puntos con textos similares se comparan a nivel de la cuarta cifra decimal, la cual corresponde a una precisión de 100 metros. Puntos en diferentes fuentes que tengan textos similares y coordenadas geográficas iguales hasta la cuarta cifra decimal se consideran el mismo punto, y se deja la fuente de mayor confiabilidad (generalmente las fuentes oficiales son más confiables respecto a las alternativas).

A continuación, se describe el proceso específico para cada una de las fuentes consideradas en las tablas 4 y 5.

2.3.2.1 Censo Nacional de Población y Vivienda 2018 (CNPV 2018)

Los registros georreferenciados del Censo Nacional de Población y Vivienda 2018 cuentan con los siguientes atributos para la selección de los puntos adecuados para la generación de las ACER:

- UA_CLASE: indica la clase a la que pertenece el punto: '1' si es cabecera municipal, '2' si es centro poblado y '3' si es rural disperso. Para el ejercicio, son de interés los puntos con UA_CLASE = '3'.
- UVA_USO_UN: atributo que indica el uso que presenta la unidad donde se aplicó el

cuestionario, con los siguientes valores: '1' si es residencial, '2' si es uso mixto (incluye usos adicionales al puramente residencial) y '3' si es no residencial. Se seleccionan los puntos cuyo UVA_USO_UN = '2' o bien UVA_USO_UN = '3' (no se incluyen usos residenciales).

- UVA1_COD_O: atributo que tiene un valor diferente de NULL si UVA_USO_UN = '2'. Son de interés los puntos con valores UVA1_COD_O = '1' (uso mixto industrial), UVA1_COD_O = '2' (uso mixto comercial) y UVA1_COD_O = '3' (uso mixto de servicios).
- UVA2_UNDNO: atributo que es distinto de NULL si UVA_USO_UN = '3' (uso no residencial). Se seleccionan los puntos con UVA2_UNDNO = '1' (industrial), UVA2_UNDNO = '2' (comercial), UVA2_UNDNO = '3' (servicios) y UVA2_UNDNO = '5' (institucional).

Lo anterior puede consolidarse obtenerse en una sola consulta SQL sobre la capa GEO_CNPV_V5_2019 (formularios georreferenciados del CNPV 2018):

SELECT * FROM GEO_CNPV_V5_2019 WHERE

- UA_CLASE = '3' AND
- ((UVA_USO_UN = '2' AND
- (UVA1_COD_O = '1' OR UVA1_COD_O = '2' OR UVA1_COD_O = '3')) OR
- (UVA_USO_UN = '3' AND
- (UVA2_UNDNO = '1' OR UVA2_UNDNO = '2' OR UVA2_UNDNO = '3' OR UVA2_UNDNO = '5'))

En síntesis, de la fuente CNPV 2018 se seleccionan aquellos puntos ubicados en el rural disperso (clase 3) cuyo uso sea mixto o no residencial, y que sean de industria, comercio, servicios o

institucionales, lo que da un total de 108.135 registros válidos.

2.3.2.2 Directorio Estadístico Nacional de Empresas – DEST

El Directorio Estadístico Nacional reúne a todas las empresas con actividad en el último año, debidamente georreferenciadas. Al excluir los puntos ubicados en cabeceras municipales y centros poblados, se tiene un total de 1.663 registros.

2.3.2.3 Capa CONSTRUCCION_P de la cartografía base 1:100.000 del IGAC

La capa CONSTRUCCION_P, presente en la cartografía base 1:100.000 del IGAC, se define en el Catálogo de Objetos de Cartografía Básica como la "obra civil ya construida o edificada, destinada a varios tipos de usos" (IGAC, 2016).

Dentro de la capa, se encuentra el atributo CUDE (CODIGO_USO_EDIFICACION), que hace referencia al uso que se da a la obra civil¹⁹. Los códigos de uso de interés económico con sus respectivas definiciones tomadas del Catálogo de Objetos, se muestran en la Tabla 4.

¹⁹ Marco Geocéntrico Nacional de Referencia. Sistema de Referencia Geocéntrico para las Américas

Tabla 4: Códigos de uso de edificación de interés económico en la capa CONSTRUCCION_P

Cude	Etiqueta	Definición ²⁰
4131	Hotel	Construcción destinada a prestar el servicio de hospedaje de manera temporal
2320	Sitio de interés	Lugar de alguna trascendencia histórica, cultural, económica o turística de interés general
4166	Seguridad	Construcción para uso de las Fuerzas Armadas con fines de seguridad, protección y asistencia a la comunidad
5620	Pozo	Excavación o perforación hecha en la tierra o mar para la extracción de líquidos o gases
5610	Tanque	Obra civil destinada a depósito de fluidos. Se captura como puntos los tanques para almacenamiento (se incluye esta construcción al estar frecuentemente asociada a plantas de tratamiento de agua potable)
5684	Molino	Construcción cuyo funcionamiento permite la transformación o extracción de elementos vitales como el viento y el agua
5608	Silo	Construcción utilizada como un depósito de alimentos, principalmente de granos.
3413	Faro	Torre alta iluminada situada en las costas o islas que sirve de guía a los navegantes durante la noche. Su ubicación aislada a lo largo de vías navegables los hace sumamente valiosos como marcas terrestres
3680	Industria	Área que contiene edificaciones e instalaciones destinadas a la actividad industrial, dotado de la maquinaria, herramienta e instalaciones necesarias para la fabricación de ciertos objetos y obtención de determinados productos
4119	Salud	Instalación o conjunto de instalaciones, empleadas para brindar servicios de prevención, diagnóstico, hospitalización y rehabilitación a pacientes (puesto o centro de salud, clínica, hospital, centros médicos, IPS)
4129	Iglesia	Edificación construida para la congregación de personas que practican determinado culto religioso

Fuente: DANE. Censo Económico

¹⁸ Marco Geocéntrico Nacional de Referencia. Sistema de Referencia Geocéntrico para las Américas

Se descartan usos como "Otras construcciones" (CUDE = 4101) al no tener certeza que, si corresponde o no a una edificación de interés económico, e "Instalación Minera" (2302), debido a que estas unidades son objeto del futuro Censo Minero.

Al seleccionar los puntos con los CUDE listados en la tabla y descartar los que se encuentran en cabeceras municipales o centros poblados, se tienen 1.287 puntos de construcciones fuente IGAC con actividad económica en el rural disperso.

2.3.2.4 Mapas de Ruta Colombia, fuente IGAC

La publicación Mapas de Ruta Colombia IGAC muestra una selección de 18 rutas turísticas entre cabeceras municipales de importancia en el país. Cada mapa de ruta consta de una Geodatabase (GDB), la que a su vez contiene la capa SitiosRutaColombia que clasifica y especifica los sitios de interés en categorías, tales como atractivos históricos, bellezas escénicas, estaciones de gasolina, hoteles, parques temáticos²¹.

Una vez se realiza la depuración de esta capa, se tienen 9.740 registros localizados en la clase 3

2.3.2.4 Mapas de Ruta Colombia, fuente IGAC

La publicación Mapas de Ruta Colombia IGAC muestra una selección de 18 rutas turísticas entre cabeceras municipales de importancia en el país. Cada mapa de ruta consta de una Geodatabase (GDB), la que a su vez contiene la capa SitiosRutaColombia que clasifica y especifica los sitios de interés en categorías, tales como atractivos históricos, bellezas escénicas, estaciones de gasolina, hoteles, parques temáticos".

Una vez se realiza la depuración de esta capa, se tienen 9.740 registros localizados en la clase 3.

2.3.2.5 Plantas de beneficio animal y vegetal, fuente INVIMA

Cabe aclarar que en esta capa no es posible georreferenciar todos los registros, en especial los rurales, al no poder vincular la dirección del establecimiento a una ubicación espacial mediante el georreferenciador automático del DANE, debido a limitaciones de la base de direcciones de la herramienta. Excluyendo los puntos de cabeceras municipales y centros poblados, así como las filas sin coordenadas (valor NULL), se llega a un total de 74 plantas de beneficio animal y vegetal.

2.3.2.6 Establecimientos hoteleros (Registro Nacional de Turismo, fuente MINCIT)

Considerando las mismas observaciones de la fuente INVIMA en cuanto a la imposibilidad de georreferenciar algunos registros y excluyendo los puntos de las cabeceras municipales y centros poblados, se obtienen 21 puntos.

Debido a la baja cantidad de registros, fue necesario recurrir a una fuente alternativa de información, teniendo en este caso los registros de la página Booking.com (para reserva de habitaciones o cupos en establecimientos hoteleros), descartando esta fuente. Booking.com se describirá en el inciso de fuentes alternativas de información.

²¹ Instituto Geográfico Agustín Codazzi (2018). Documento técnico de mapas de ruta Colombia.

2.3.2.7 Peajes (fuente Agencia Nacional de Infraestructura ANI)

La publicación Mapas de Ruta Colombia IGAC muestra una selección de 18 rutas turísticas entre cabeceras municipales de importancia en el país. Cada mapa de ruta consta de una Geodatabase (GDB), la que a su vez contiene la capa SitiosRutaColombia que clasifica y especifica los sitios de interés en categorías, tales como atractivos históricos, bellezas escénicas, estaciones de gasolina, hoteles, parques temáticos²¹.

Una vez se realiza la depuración de esta capa, se tienen 9.740 registros localizados en la clase 3

2.3.2.8 Instituciones Prestadoras de Salud (IPS, fuente MINSALUD)

Excluyendo aquellos establecimientos presentes en clases 1 y 2, la base de datos de IPS (instituciones prestadoras de salud) reportadas por el Ministerio de Salud, tienen 3.880 registros, todos georreferenciados.

2.3.2.9 Establecimientos educativos (fuente MEN)

Los establecimientos educativos incluyen instituciones de preescolar, primaria y secundaria (no incluyen instituciones de educación superior). En el rural disperso, se encuentran 26.931 puntos.

2.3.2.10 Open Street Map (fuente alternativa)

Al descargar toda la base de datos correspondiente a Colombia de la dirección <https://download.geofabrik.de/south-america/colombia.html>, seleccionar la capa Points of Interest (POI) y realizar la selección por localización para extraer los puntos ubicados en la zona rural, se obtiene una capa con 9.034 registros.

Al igual que la capa de construcciones del IGAC, la capa Points of Interest maneja una codificación de los usos de interés económico, la cual se puede

consultar en la página <https://www.geofabrik.de/data/geofabrik-osm-gis-standard-0.6.pdf>.

2.3.2.11 Booking.com

La página <https://www.booking.com/> ofrece el servicio de reserva y alquiler de alojamientos a nivel mundial, en establecimientos de diversa categoría (hoteles, hostales, posadas, fincas, etc.). Con el fin de realizar una selección informada del alojamiento, se incluyen datos relevantes del lugar, incluyendo sus coordenadas geográficas en muchos casos. Mediante técnicas de web scraping es posible almacenar los registros georreferenciados de este sitio web para Colombia, llegando a un total de 2.509 puntos.

2.3.2.12 Integración de los conjuntos de datos en una sola capa

Antes de realizar la integración de todos los conjuntos de datos, es preciso señalar que los duplicados en la misma fuente no se eliminan, debido a que en algunas capas (en especial DEST), la misma coordenada geográfica corresponde a diferentes establecimientos económicos.

De cualquier modo, la depuración de las fuentes asegura la calidad del resultado final de la aplicación de la metodología, por lo que este proceso está sujeto a constante revisión y actualización.

El resultado de la consolidación de las fuentes que se encontraban en capas separadas, se presenta en la Tabla 5.

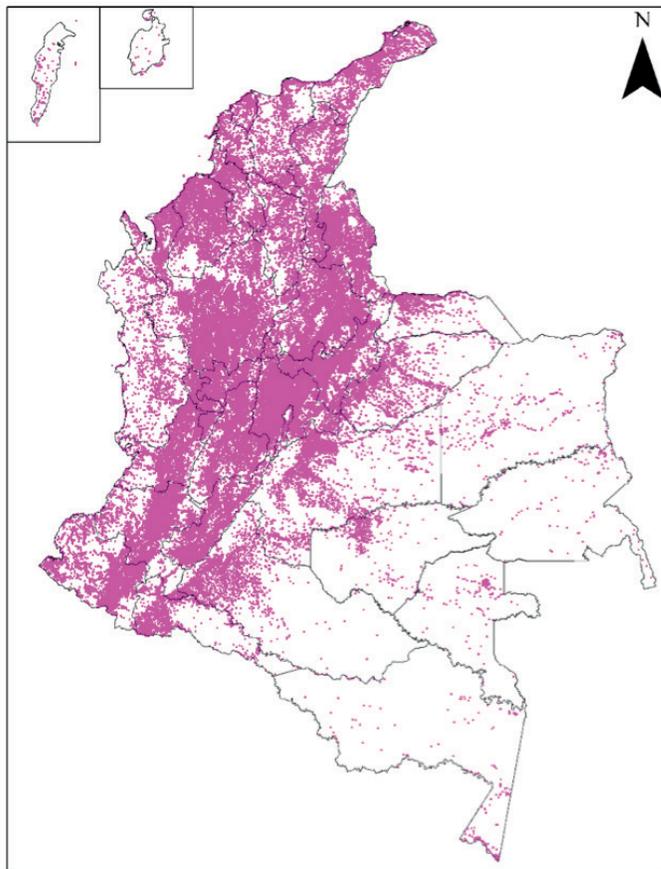
Tabla 5: Códigos de uso de edificación de interés económico en la capa CONSTRUCCION_P

CAPA	NOMBRE	NÚMERO DE REGISTROS
CNPV_2018	Censo Nacional de Población y Vivienda 2018	108.135
DEST_2020	Directorio Estadístico Nacional DANE 2020	1.663
CONSTRUCCION_P_IGAC_2017	Construcción IGAC 2017	1.287
MAPAS_RUTA_COLOMBIA_2018	Mapas de Ruta Colombia IGAC 2018	9.740
INVIMA_2019	Plantas de Beneficio Animal y Vegetal 2019	74
PEAJES_ANI_2020	Peajes INVIAS 2020	99
EST_EDUCATIVOS_2019	Establecimientos Educativos MEN 2019	26.931
IPS_2019	Instituciones Prestadoras de Salud (IPS) MINSALUD 2019	3.880
BOOKING.COM	Base de alojamientos BOOKING 2020	2.509
OSM_2020	Puntos de Interés Open Street Map 2020	9.034
TOTAL		163.445

Fuente: DANE. Censo Económico

Para efectos de esta metodología, todas las capas se unifican en una sola entidad geográfica, denominada GEO_FUENTES_CONSOLIDADAS_2020, en sistema de coordenadas geográficas MAGNA. Este es el insumo principal para el flujo de procesos geoespaciales que genera los polígonos de las ACER.

Figura 2. Distribución geográfica de la capa consolidada de fuentes de información GEO_FUENTES_CONSOLIDADAS_2020



Fuente: DANE. Censo Económico

En la figura 2, se observa una tendencia de concentración de establecimientos económicos en las regiones Caribe y Andina, al igual que hacia el piedemonte de la Cordillera Oriental. Hay escasez de información en el Andén Pacífico y en las regiones Orinoquía y Amazonía.

La fuente que aporta un mayor número de registros de establecimientos económicos es el CNPV 2018.

2.3.3 Selección de puntos por municipio y reproyección de capas a coordenadas planas

Una vez consolidada la capa compuesta de todas las fuentes (GEO_FUENTES_CONSOLIDADAS_2020), se procede a la selección de los puntos que se encuentran en el municipio que se está evaluando, para ir a continuación a reproyectar esta capa en las coordenadas planas (origen MAGNA SIRGAS al que pertenece el municipio y que ya se ha determinado previamente).

La secuencia de etapas en ArcGIS ModelBuilder es el siguiente:

- Creación de la File Geodatabase o GDB donde se almacenarán todas las capas del municipio, con el nombre CE_ACER_XXXXX (donde XXXXX es el código DIVIPOLA del municipio).
- Selección del polígono del municipio mediante el atributo MPIO_CCNCTNDO en la capa CNPV_MPIO_ORIGEN (capa DANE de límites municipales, a la que se le ha adicionado el atributo ORIGEN con el sistema de coordenadas planas MAGNA SIRGAS por municipio).
- Selección por localización de los puntos de la capa GEO_FUENTES_CONSOLIDADAS_2020 que se encuentran dentro del polígono municipal.
- Verificación de que todos los puntos pertenecen al municipio dado mediante comparación del valor de MPIO_CCNCTNDO con los valores del atributo MPIO. Puntos con valores distintos al del municipio seleccionado se eliminan con la herramienta Delete Features.
- Reproyección al sistema de coordenadas planas MAGNA SIRGAS de los puntos seleccionados y del límite municipal, almacenando las capas creadas en "CE_ACER_XXXXX.gdb".

El resultado final de este proceso es la capa de los establecimientos económicos de un municipio en coordenadas planas en alguno de los 6 orígenes MAGNA SIRGAS (dependiendo del municipio), y el polígono cuyo contorno es el límite del municipio, también en coordenadas planas. Estos insumos entran a la siguiente etapa, que consiste en el cálculo de la métrica estadística adecuada para la generación de los polígonos de referencia de las ACER.

2.3.4 Cálculo de la mediana de las distancias de los establecimientos a su vecino más cercano

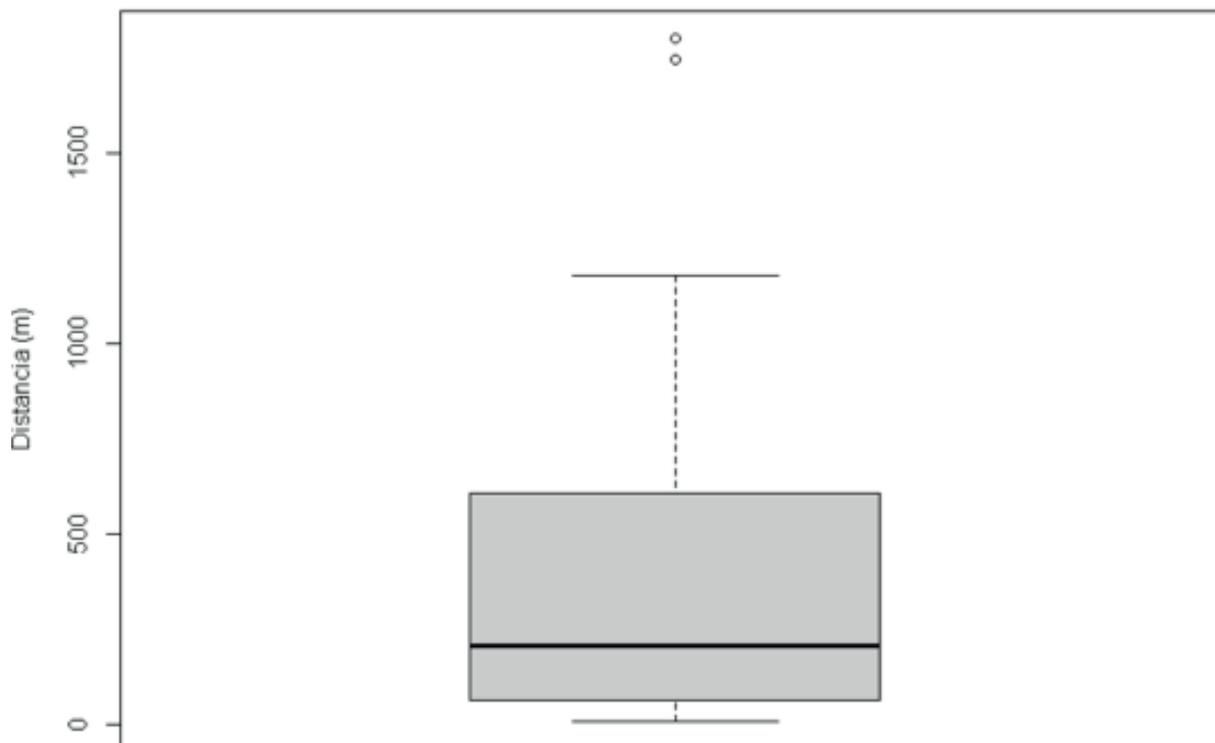
Dentro del ejercicio de determinación de ACER, se busca establecer un parámetro que, por una parte, pueda caracterizar a un municipio dado, y por otra, ser de utilidad para realizar cálculos posteriores, sin intervención o subjetividad del investigador que ejecute el proceso.

Herramientas de sistemas de información geográfica como ArcMap® o QGIS® emplean el promedio de distancias al vecino más cercano como un índice estándar para establecer si un conjunto puntual tiene una distribución aleatoria en el espacio, o por el contrario existe algún grado de concentración, y también como una medida típica de los datos dentro de un área determinada.

Sin embargo, en las áreas rurales de los municipios de Colombia, los datos de distancias a vecinos más cercanos tienen una alta probabilidad de incluir datos atípicos, debido a que se consideran fuentes secundarias de diferente origen y vigencia. También se encuentran puntos muy alejados unos de otros por las características propias de los municipios (presencia de predios de gran extensión, dificultades en vías de comunicación, etc.). En estas condiciones, el promedio no representa la tendencia central de la variable de interés, en este caso la distancia de un establecimiento económico a su vecino más cercano.

En la figura 3 se observa el diagrama de caja y bigotes de las distancias al vecino más cercano de los establecimientos económicos del municipio de Ubaté (Cundinamarca), donde se aprecian dos valores atípicos a más de 1500 m, lo cual podría afectar el promedio aritmético.

Figura 3. Diagrama de caja y bigotes de las distancias de los establecimientos económicos a su vecino más cercano para el municipio de Ubaté (Cundinamarca)



Fuente: DANE. Censo Económico

Se ha propuesto emplear la mediana de las distancias al vecino más cercano en reemplazo de la media aritmética, con el fin de establecer una medida estadística más robusta para cada municipio sin el sesgo que implican los valores atípicos, como los observados en el caso de Ubaté. Con este criterio, se ejecutan los siguientes pasos:

- Se crea una copia de la capa de establecimientos económicos por municipio y esta se depura, eliminando los puntos con iguales coordenadas (longitud y latitud). Este paso es necesario para evitar el cálculo de distancias iguales a cero, que afectan el valor final de la mediana.
- Se emplea la herramienta Near de ArcGIS para calcular las distancias de cada punto a su vecino más cercano.
- Con la tabla resultante, se ejecuta un script en el lenguaje de programación Python dentro de la herramienta Calculate Value, la cual determina la mediana de las distancias de los establecimientos a su vecino más cercano (ArcGIS no tiene una herramienta que la calcule propiamente).
- El valor resultante se emplea en los cálculos posteriores, como una distancia central característica y única para cada municipio.

2.3.5 Geoprocesos de generación de polígonos de referencia de las ACER

Una vez calculada la mediana de las distancias al vecino más cercano, denotada de ahora en adelante como **Me**, la capa original de establecimientos económicos (sin haber eliminado los registros repetidos por localización) entra al proceso de generación de los polígonos de referencia de las ACER, el cual consta de las siguientes etapas:

Aplicación de algoritmos para hallar áreas de concentración de establecimientos económicos

La capa original de establecimientos económicos, junto con el límite del municipio y el valor de **Me**, entran a dos geoprocesos paralelos: **índices locales de autocorrelación espacial (LISA)** por sus siglas en inglés) y **área de influencia (Buffer)** por el nombre de la herramienta original).

Procesamiento de los polígonos resultantes de los algoritmos

Los polígonos resultantes de los procesos anteriores deben pasar por otros geoprocesos para obtener finalmente los polígonos de referencia de las ACER del municipio. Incluyen herramientas como disolución, vínculo espacial y selección por atributos, los cuales se explicarán más adelante.

2.3.5.1 Índices locales de autocorrelación espacial (LISA)

Los índices locales de autocorrelación espacial (LISA), son medidas de significancia estadística que permiten representar si un fenómeno localizado

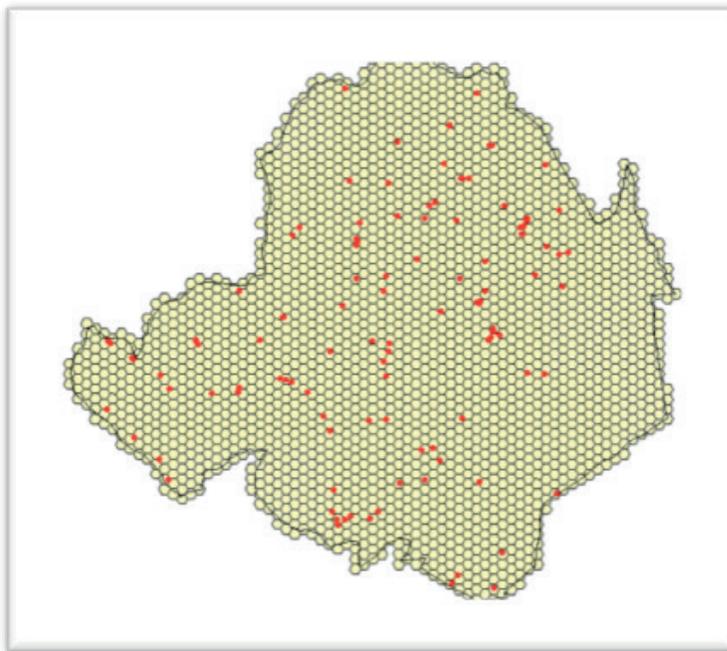
espacialmente conforma aglomeraciones o clústers, si existen outliers o valores dispersos, o si su distribución es aleatoria. El algoritmo se basa en el cálculo de los índices locales de Moran, I, así como en los valores z (desviaciones estándar) y valores p (probabilidades). Para mayor información, consultar la referencia²².

Para hallar las áreas donde se concentran los establecimientos económicos mediante este algoritmo, se siguen los pasos descritos a continuación:

- Creación de una grilla de agregación. El análisis de los eventos (en este caso los establecimientos económicos) no se realizará por sí solos, sino agregados en una grilla que se extiende por toda el área del municipio. Esta grilla presenta las siguientes particularidades:
 1. Es hexagonal, es decir, formada por celdas en forma de hexágono regular. Un hexágono tiene en total 6 celdas contiguas.
 2. La altura de cada celda es igual a **Me**, es decir, las dimensiones de la celda de la grilla son únicas para cada municipio por cuanto **Me** varía de acuerdo con la distribución de los establecimientos.
- De esta forma, se crea una grilla con celdas hexagonales, cuya altura (la distancia entre dos aristas paralelas del hexágono) es igual a la mediana de las distancias de los establecimientos económicos a su vecino más cercano, lo cual asegura la generación de una grilla exclusiva de cada municipio y cuyas dimensiones están directamente relacionadas con la distribución espacial de los establecimientos.

²² Metodología ESRI (2020). Tomado de: <https://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-statistics-toolbox/h-how-cluster-and-outlier-analysis-anselin-local-m.htm>

Figura 4. Grilla hexagonal para el municipio de Ubaté (Cundinamarca) y establecimientos económicos, donde la altura del hexágono viene dada por $Me = 207,7659$ m



Fuente: DANE. Censo Económico

- **Conteo de establecimientos económicos por cada celda de la grilla.** El siguiente paso consiste en realizar un conteo de establecimientos económicos (puntos) por cada hexágono o celda de la grilla, proceso conocido como agregación. El resultado es la misma grilla, con un atributo o columna que indica cuántos establecimientos hay en cada celda o hexágono.
- **De esta forma,** se logra pasar de un fenómeno puntual (los establecimientos en sí mismos) a un fenómeno de áreas (las celdas o hexágonos) donde cada celda tiene un valor particular de establecimientos económicos, mayor o igual a cero. Es sobre estas celdas sobre las que se calcula el índice de autocorrelación espacial o LISA.
- **Cálculo de LISA.** La grilla resultante de la operación anterior es la capa de entrada de la herramienta LISA, la cual requiere una variable de interés. Esta corresponde al número de establecimientos en cada celda. Al realizar los cálculos de los índices locales de Moran y de los valores z y p mencionados anteriormente, interesan aquellos hexágonos con valores estadísticamente significativos de clúster ALTO-ALTO (celdas con valores altos de establecimientos rodeadas de otros valores altos) y outlier ALTO-BAJO (celdas con valores altos rodeadas de valores bajos).
- **Selección de hexágonos de interés.** El algoritmo genera una nueva grilla con un atributo que puede tener cuatro valores: "HH" (alto - alto), "HL" (alto - bajo), "LH" (bajo - alto) y "LL" (bajo - bajo). Se seleccionan las celdas que tengan valores "HH" y "HL" y se separan para conformar una nueva capa correspondiente a la salida de este algoritmo.

2.3.5.2 Área de influencia (Buffer)

Consiste en la generación de áreas de influencia circulares (buffers, en inglés) sobre los establecimientos económicos, fusionando las áreas resultantes de manera que formen un solo polígono.

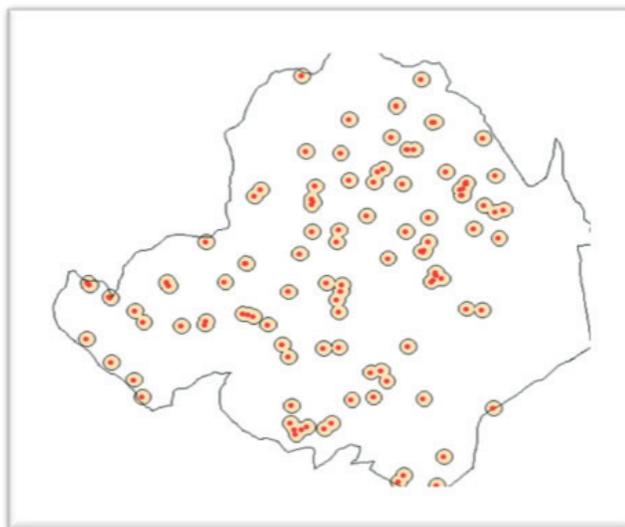
Este proceso se basa en considerar que un establecimiento económico posee un área de influencia circular. Esta área no es infinita, y se considera que su límite es el área de influencia de otro establecimiento vecino.

Siendo la distancia **Me** una medida representativa del municipio, se propone que el área de influencia de un establecimiento se extiende máximo a una

distancia igual o menor a **Me**. Entre más cercanos se encuentren un grupo de establecimientos –esto es, más concentrados –, sus áreas de influencia tienden a solaparse y a fusionarse en una sola área que funciona como zona de influencia de todos los establecimientos. Dicho de otra manera, una concentración de establecimientos generará un área de influencia común a todos ellos, un área de concentración económica rural o ACER.

El algoritmo consta de un solo paso, que consiste en la generación de círculos alrededor de los puntos con un radio igual a **Me** y la fusión de los círculos que se traslapan.

Figura 5. Áreas de influencia (Buffer) de los establecimientos económicos del municipio de Ubaté (Cundinamarca). El radio del buffer corresponde a $Me = 207,7659$ m



Fuente: DANE. Censo Económico

2.3.5.3 Procesamiento de los polígonos resultantes de los algoritmos

En cada uno de los algoritmos anteriores se generan polígonos que originalmente están separados y no conforman geoméricamente un área consolidada. Por tal razón es necesario aplicar operaciones geoespaciales adicionales para generar los polígonos de referencia de las ACER.

El procedimiento es el siguiente:

- Los polígonos generados (hexágonos o buffers) se fusionan creando un nuevo polígono, haciendo que polígonos no conexos sean independientes (no se forma una entidad multiparte).

- Mediante la operación de vínculo espacial, se determina cuántos establecimientos económicos (puntos) se encuentran dentro de cada polígono resultante de la fusión (un procedimiento similar al conteo de establecimientos dentro de los hexágonos, que crea un nuevo atributo con el conteo de establecimientos por polígono).
- Con una operación de selección por atributos, **se escogen los polígonos que tengan un número mayor o igual a la mínima cantidad de establecimientos económicos que debe tener un ACER**. Este parámetro viene definido por el usuario al momento de correr el geoproceso en su totalidad y hasta el momento constituye la única variable dentro del modelo que define el investigador.
- El resultado final son los polígonos de referencia de las ACER para cada uno de los algoritmos mencionados.

2.3.5.4 Obtención de los polígonos de referencia de las ACER

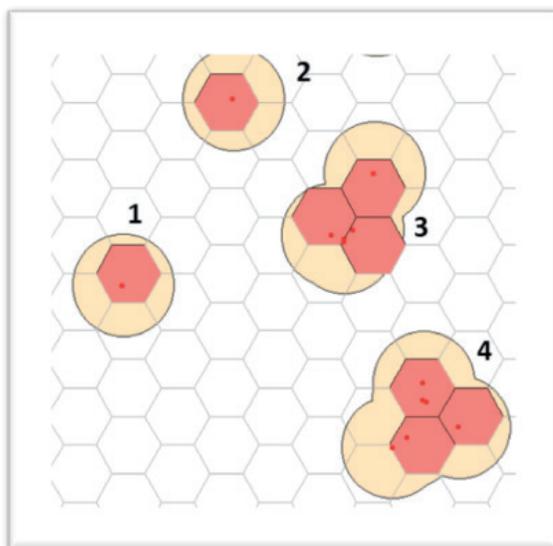
El siguiente paso consiste en obtener los polígonos de referencia de las ACER a partir de los resultados anteriores.

En aras de asegurar la existencia de concentración de actividades económicas en el área rural, se busca garantizar la simultaneidad espacial (que ocurran en el mismo sitio) de los siguientes fenómenos:

- Evidencia de concentración espacial de establecimientos económicos, confirmada por significancia estadística (LISA).
- Evidencia de áreas de influencia comunes para varios establecimientos económicos cercanos (Buffer).

De este modo, los polígonos de referencia de las ACER serán el resultado de determinar los polígonos de buffer que se intersecten con los polígonos de LISA, donde se encuentre un número mayor o igual al mínimo de establecimientos económicos.

Figura 6. Superposición de polígonos LISA y Buffer en el municipio de Ubaté (Cundinamarca) obtenidos sobre los establecimientos económicos



Fuente: DANE. Censo Económico

En la figura 6, las intersecciones identificadas como 1 y 2 se descartan (tienen un solo establecimiento cada una), mientras que se mantienen las intersecciones 3 y 4 (presentan más de 3 establecimientos).

En síntesis, el polígono de referencia del ACER será el área de influencia que a su vez contenga un polígono LISA con más de 3 establecimientos económicos.

2.3.6 Validación de ACER mediante variables exógenas

Como se indicó en el inciso de alcance de las ACER, la identificación y delimitación de los polígonos de referencia se basa únicamente en criterios geoespaciales y en variables endógenas (la localización de los puntos y sus medidas de tendencia central), sin tener en cuenta las variables socioeconómicas de los establecimientos, al carecer de información suficiente en la mayoría de las fuentes.

Por lo anterior, no es posible determinar a priori si los polígonos de referencia de ACER calculados son representativos de la actividad económica en el rural disperso de su municipio respectivo. Este hecho resulta de gran importancia para las dependencias del DANE encargadas de la logística del Censo Económico 2021, dado que el resultado de la metodología debe asegurar que en las áreas demarcadas como ACER se encuentren establecimientos económicos para no incurrir en pérdida de tiempo y recursos desplazándose a zonas no representativas.

Una manera de priorizar las ACER es construyendo un área potencial o un área óptima de actividad económica en el rural disperso de cada municipio, empleando variables relacionadas con el territorio y con la dinámica económica, y posteriormente cruzando dicha área con las ACER para seleccionar aquellas que se encuentren en la intersección. A las variables empleadas para construir el área óptima se les conoce como exógenas, ya que son externas a la ubicación de los establecimientos económicos.

El procedimiento seleccionado para la determinación del área óptima es el análisis o evaluación multicriterio o jerárquico (conocido por sus siglas en inglés, AHP), el cual se explica a continuación.

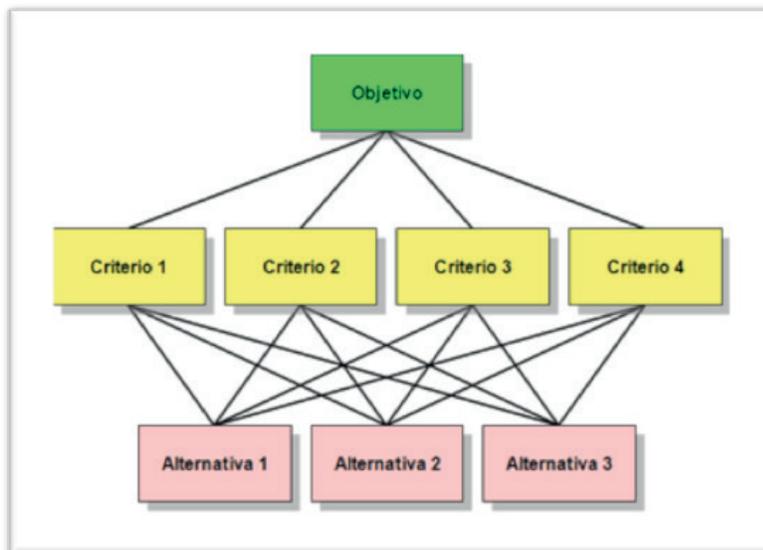
2.3.6.1 Definición de la evaluación multicriterio

La evaluación multicriterio, también conocida como análisis jerárquico, es una técnica desarrollada por T. Saaty en 1980²³, donde inicialmente se establece un objetivo y unas ciertas alternativas para cumplirlo confrontadas con ciertos criterios, los cuales se modelan a través de pesos o porcentajes que califican y evalúan las alternativas, para escoger alguna o para efectuar una suma ponderada de los criterios.

Esta metodología busca justificar una decisión subjetiva, donde los juicios sobre las alternativas se transforman en análisis cuantitativos empleando escalas de importancia relativa, lo que permite dar soluciones coherentes y sustentadas.

²³ Saaty, T.L (1980). The Analytics Hierarchy Process: Planning, Priority Setting, Resource Allocation, Mc. Graw Hill.

Figura 7. Esquema general de la evaluación multicriterio.



Fuente: <http://www.medicalsapiens.cl/AnMulticriterio.php>

Los criterios empleados en este tipo de análisis deben cumplir las siguientes características:

- No ser redundantes, dicho de otra forma, ser independientes o mutuamente ortogonales entre sí.
- En lo posible, deben dividirse en subcriterios para simplificar el proceso.
- Deben ser precisos y consistentes con el objetivo.
- En cada evaluación, se recomienda definir máximo siete (7) criterios.

2.3.6.2 Pasos de la evaluación multicriterio

El siguiente es un esquema general de los pasos o etapas de la evaluación multicriterio.

- Definición del problema planteado y de la solución deseada al mismo.
- Estructurar jerarquías de criterios y alternativas: definición de variables y rangos.
- Construir la matriz de comparación por pares de criterios (MCP). Esta es una matriz en donde los criterios o variables se disponen tanto en filas como en columnas y se comparan dos a la vez (calificación pareada), empleando unos puntajes específicos.
- Cálculo del vector de pesos de los criterios mediante operaciones de álgebra matricial.
- Comprobación de la consistencia de la matriz de comparación por pares, mediante el cálculo de la razón de consistencia (RC).
- Asignación de los pesos obtenidos a cada criterio y cálculo de su suma ponderada.

2.3.6.3 Definición del problema y de la solución mediante variables exógenas

En primer lugar, se establece el problema a tratar con el análisis multicriterio, que consiste en conocer cuáles ACER representan la actividad económica en un municipio determinado, empleando variables exógenas. Su solución está dada por el cálculo de un área óptima o potencial de actividad económica en el rural disperso del municipio a partir de la suma ponderada de las variables exógenas que se definen a continuación:

Proximidad a las vías terrestres: Distancia al eje de una vía terrestre, teniendo en cuenta la topografía del terreno.

Intensidad de luces nocturnas: Flujo radiante por unidad de área (radiancia), proveniente de las fuentes de luz nocturnas de la superficie terrestre.

Pendiente del terreno: Inclinación del terreno, medida en porcentaje.

Densidad de población: Número de personas por unidad de área.

Proximidad a cuerpos de agua: Distancia al eje de una vía fluvial de más de 50 metros de ancho de cauce, o a la orilla de cuerpos de agua como lagos, lagunas, embalses, jagüeyes, etc., teniendo en cuenta la topografía del terreno.

La selección de estas variables particulares se justifica en las siguientes razones:

- La información de las variables seleccionadas está disponible para todo el territorio nacional.
- Son variables relacionadas con la actividad económica en el rural disperso de los municipios colombianos, de acuerdo con el criterio de los expertos temáticos.
- Algunas de estas variables, como las luces nocturnas, son proxy de actividad económica.
- Las variables son independientes entre sí y de la ubicación de establecimientos económicos.



Desde el punto de vista computacional, estas variables corresponden a archivos tipo ráster o imágenes geolocalizadas, cada una de las cuales se multiplica por el porcentaje obtenido en la evaluación multicriterio y finalmente obtener el área óptima o potencial de actividad económica en el rural disperso.

2.3.6.4 Definición de rangos de las variables

Las variables definidas en el literal anterior deben clasificarse en rangos definidos por un número de clases. De esta manera, se establecen puntajes dentro de las propias variables, que permiten su calificación posterior.

Para todas las variables exógenas se estableció un total de cinco (5) clases, por ser este el número mínimo recomendado por la literatura (debe estar entre 5 y 15 clases).

A continuación, se describe la forma en que se establecieron los rangos para cada variable.

2.3.6.4.1 Distancia a vías terrestres

En otros trabajos que incluyen análisis multicriterio, los rangos de distancias a las vías dependen del tipo de estudio. Por tal razón, en este ejercicio se determinaron los valores de cada clase tomando los quintiles de distribución de distancias de los establecimientos económicos a las vías.

Teniendo en cuenta que los tipos de vía más utilizados en las áreas rurales no son los mismos para todos los municipios y varían de acuerdo con la región natural donde se encuentren. En primer lugar, se establecieron los tipos de vía definidos por el IGAC predominantes en cada región comparando la capa Vía de la cartografía base 1:100.000 del IGAC con la ubicación de los polígonos de referencia de las ACER y verificando con imágenes de satélite, los cuales se listan en la tabla 6.

Tabla 6: Total de registros georreferenciados (puntos) por fuente de información

Región natural	Tipos de vía predominantes en el rural disperso
Amazonia	1,2,3,4,5
Andina	1,2,3
Caribe	1,2,3,4
Orinoquia	1,2,3,4
Pacífica	1,2,3,4,5

Fuente: DANE. Censo Económico

Los tipos de vía a los que hace referencia la tabla 5 se muestran en la figura 8.

Figura 8. Tipos de vía definidos en la cartografía base del IGAC, escala 1:100.000

Subtipo - TIPO_VIA			
Etiqueta	Código	Definición	Imagen
Vía Tipo 1	1	Vía pavimentada, más de dos carriles para el tráfico vehicular, animal o personas, cuyo ancho excede de 5m.	
Vía Tipo 2	2	Vía Sin Pavimentar, más de dos carriles, para el tráfico vehicular, animal o personas, cuyo ancho excede de 5m.	
Vía Tipo 3	3	Vía Pavimentada y angosta, Transitible todo el año, cuyo ancho es inferior o igual a 5m.	
Vía Tipo 4	4	Vía Sin Pavimentar y angosta, para el tráfico vehicular, animal o personas, cuyo ancho es inferior o igual a 5m.	
Vía Tipo 5	5	Carreteable transitable en tiempo seco, para el tráfico vehicular, animal o personas.	

Fuente: https://www.igac.gov.co/sites/igac.gov.co/files/anexo_1.1_catalogo_objetos_cartografiabasica_v1.0_.pdf

A continuación, se estableció la distribución de distancias de los establecimientos económicos a las vías, diferenciando por región. De esta forma, los establecimientos de los 1121 municipios se clasificaron en su respectiva región natural y se calculó la distancia de cada punto a la vía más próxima, obteniendo una distribución distinta en cada región.

En la tabla 7 se muestra la distribución de distancias de los establecimientos económicos a la vía más cercana para los municipios de las 5 regiones donde predominan las vías tipo 1 hasta la 5.

Tabla 7. Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3,4,5

Quintil	Distancias en metros				
	Andina	Caribe	Pacífica	Orinoquía	Amazonía
20%	35	72	1017	55	97
40%	106	293	5218	172	754
60%	301	957	17469	637	2794
80%	867	2627	42344	2410	9946

Fuente: DANE. Censo Económico

En la tabla 8 se muestra la distribución de distancias para los municipios donde predominan las vías tipo 1 hasta la 4 (principalmente en el Caribe y la Orinoquía).

Tabla 8. Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3,4

Quintil	Distancias en metros	
	Caribe	Pacífica
20%	196	107
40%	995	724
60%	2664	2593
80%	6584	10665

Fuente: DANE. Censo Económico

En la tabla 9 se muestra la distribución que predomina en la región Andina, donde se encuentra la mayor densidad de vías y en los municipios se emplean las vías tipo 1, 2 y 3.

Tabla 9. Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3

Quintil	Andina (m)
0% a 20%	159
21% a 40%	955
41% a 60%	2513
61% a 80%	5503
81% a 100%	3.000.000

Fuente: DANE. Censo Económico

Las distribuciones anteriores permiten crear cinco clases, cuyo valor mínimo es cero y el valor máximo se fija en 3'000.000.000 de metros para todas las regiones.

2.3.6.4.2 Proximidad a cuerpos de agua

De la misma manera que se establecieron los rangos para la proximidad a vías terrestres, se definen los rangos para la proximidad a cuerpos de agua. También se lleva a cabo la diferenciación entre regiones y se calcula la distancia de cada establecimiento económico al cuerpo de agua más cercano, para establecer la distribución por región.

Los cuerpos de agua no incluyen los llamados drenajes sencillos de la cartografía base del IGAC, puesto que estos se encuentran en gran número en el rural disperso y en muchas ocasiones no corresponden a vías navegables o aprovechables.

Los resultados de la distribución de distancias por región se muestran en la tabla 10.

Tabla 10. Distribución de distancias de establecimientos económicos a la vía más cercana según región, tomando tipos de vía 1,2,3,4,5

Quintil	Distancias en metros				
	Andina	Caribe	Pacífica	Orinoquía	Amazonía
0% a 20%	1487	731	104	633	515
21% a 40%	3666	2360	356	1482	1810
41% a 60%	6288	4702	1447	2794	4296
61% a 80%	9828	8469	4648	5253	8424
81% a 100%	3.000.000	3.000.000	3.000.000	3.000.000	3.000.000

Fuente: DANE. Censo Económico

De la misma forma que para la proximidad a vías terrestres, los rangos se establecen con un valor mínimo de cero y un máximo de 3.000.000 de metros.

2.3.6.4.3 Luces nocturnas

Las investigaciones sobre los rangos de luces nocturnas a nivel internacional se han orientado a diferenciar las zonas urbanas de las rurales, sin definir rangos o clases dentro de estas últimas. Por otra parte, los valores de radiancia de luz nocturna varían entre los diferentes países y regiones dentro de estos, por lo cual la determinación de los rangos se debe realizar de manera separada para el caso colombiano.

Una aproximación empírica válida para la determinación de los rangos de una variable continua como la intensidad de luces nocturnas es el empleo de un intervalo geométrico, metodología que calcula los intervalos o límites entre las clases con una progresión geométrica:

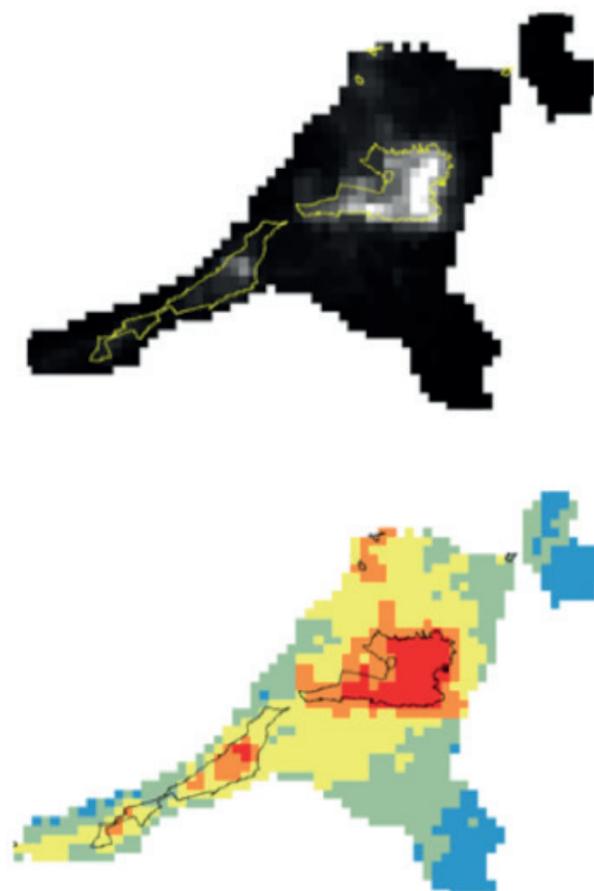
$$a_n = ar^{n-1}$$
$$r = \left(\frac{\text{valor máximo}}{\text{valor mínimo}} \right)^{1/n}$$

En la ecuación anterior, n hace referencia al número de clases deseadas, a es el valor mínimo o inicial de la progresión y r es la razón o cociente de la progresión, con el que se calculan los valores de los intervalos de clase.

Los intervalos geométricos son adecuados al permitir la clasificación de variables continuas con presencia de atípicos o valores extremos. Por otra parte, en el caso particular de las luces nocturnas, permiten diferenciar claramente las zonas urbanas de las rurales y establecer rangos de valores en estas últimas.

La figura 9 muestra la reclasificación de la imagen ráster de luces nocturnas del municipio de Fusagasugá por el método del intervalo geométrico, resaltando los límites de la cabecera municipal y de los centros poblados.

Figura 9. Reclasificación de luces nocturnas del municipio de Fusagasugá por medio de intervalo geométrico



Fuente: DANE. Censo Económico

Como las luces nocturnas varían de un lugar a otro los rangos de valores también varían, por lo que es necesario calcular el intervalo geométrico por cada municipio. De esta forma no existe un único rango de valores de luces nocturnas para todo el país, estableciendo 5 clases, las cuales quedan definidas de forma única para cada municipio dependiendo del rango de valores respectivos.

2.3.6.4.4 Densidad de población

La densidad de población en el rural disperso de cada municipio se calcula como un ráster de densidad kernel (mapa de concentración o "mapa de calor") a partir de los datos del CNPV 2018. La densidad de población rural varía de municipio a municipio y presenta valores muy altos en ciertos lugares y muy bajos en otros. Por tal razón también se reclasifica empleando un intervalo geométrico, que como ya se indicó es adecuado para este tipo de variables continuas, con valores extremos o atípicos.

Así como las luces nocturnas, los intervalos de densidad poblacional son únicos para cada municipio, estableciendo 5 clases, las cuales quedan definidas de forma única para cada municipio dependiendo del rango de valores respectivos.

2.3.6.4.5 Pendiente

A diferencia de las variables anteriores, la pendiente del terreno ha sido clasificada por diversas entidades públicas y privadas a nivel nacional. Para este ejercicio, se toma la clasificación propuesta por la Autoridad Nacional de Licencias Ambientales (ANLA), adoptada por otras entidades como el IGAC. En la tabla 11 se muestran los valores de los rangos, dados en porcentaje.

Tabla 11. Rangos de valores de pendiente

Clase	Nombre	Valor
1	PLANA O A NIVEL	0 a 3%
2	LIGERAMENTE INCLINADA	3 a 7%
3	MEDIANAMENTE INCLINADA	7 a 12%
4	FUERTEMENTE INCLINADA	12 a 25%
5	ESCARPADA O EMPINADA	Mayor a 25%

Fuente: DANE. Censo Económico

2.3.6.5 Matriz de comparación por pares (MCP)

La matriz de comparación por pares (en adelante MCP) es una matriz cuadrada, donde las variables involucradas se organizan en la primera fila y en la primera columna para luego realizar comparaciones pareadas entre ellas (Figura 10).

Figura 10. Ejemplo de matriz de comparación por pares (MCP)

MCP	Proximidad a vías terrestres (PVT)	Intensidad de Luces Nocturnas (LN)	Pendiente del Terreno (P)	Densidad de Población (DP)	Proximidad a cuerpos de agua (PVF)
Proximidad a vías terrestres (PVT)	1				
Intensidad de Luces Nocturnas (LN)		1			
Pendiente del Terreno (P)			1		
Densidad de Población (DP)				1	
Proximidad a cuerpos de agua (PVF)					1

Fuente: DANE. Censo Económico

La MCP presenta las siguientes propiedades²⁴:

Reciprocidad. Siendo a un elemento de la matriz en la fila i y la columna j . Si $a_{ij} = x$, $a_{ji} = 1/x$

Homogeneidad. Si i y j son igualmente importantes, $a_{ij} = a_{ji} = 1$ para todo i, j .

Consistencia. La matriz no debe contener contradicciones en la valoración realizada.

Las comparaciones pareadas entre variables se realizan empleando la escala fundamental definida en la figura 11.

Figura 11. Escala fundamental de comparación por pares

VALOR	DEFINICIÓN	COMENTARIOS
1	Igual importancia	El criterio A es igual de importante que el criterio B
3	Importancia moderada	La experiencia y el juicio favorecen ligeramente al criterio A sobre el B
5	Importancia grande	La experiencia y el juicio favorecen fuertemente el criterio A sobre el B
7	Importancia muy grande	El criterio A es mucho más importante que el B
9	Importancia extrema	La mayor importancia del criterio A sobre el B está fuera de toda duda
2,4,6 y 8	Valores intermedios entre los anteriores, cuando es necesario matizar	

Fuente: <https://victoryepes.blogs.upv.es/tag/saaty/>

²⁴ Yepes, V. (2018). Proceso Analítico Jerárquico (Analytic Hierarchy Process, AHP). Universitat Politècnica de València. Recuperado de <https://victoryepes.blogs.upv.es/tag/saaty/>

El procedimiento de comparación es el siguiente:

- En primer lugar, se organizan las variables de acuerdo con su importancia relativa de mayor a menor y de esta manera se disponen en la primera fila y en la primera columna de la matriz.
- A continuación, se realizan las comparaciones entre las variables de acuerdo con la figura 7. Generalmente se diligencia solo la matriz triangular superior (celdas a la derecha de la diagonal de unos), ya que los valores de la matriz triangular inferior corresponden al

inverso multiplicativo de la superior (principio de reciprocidad).

- La comparación debe hacerse por mínimo 10 expertos en matrices diferentes, ya que al tratarse de un juicio subjetivo se quieren tener varias opiniones. Las matrices resultantes se promedian utilizando una media geométrica la cual, a diferencia de la media aritmética, mantiene el principio de reciprocidad al conservar los valores a_{ij} como los inversos multiplicativos de los elementos a_{ji} .

2.3.6.6 Cálculo del vector de pesos

Una vez se tiene la matriz de comparación, se realizan las siguientes operaciones:

- Se obtiene la suma de los elementos por columna de la MCP. En la figura 12 se muestra una matriz de ejemplo con este cálculo.

Figura 12. Suma por columna de la MCP

MCP	Proximidad a vías terrestres (PVT)	Intensidad de Luces Nocturnas (LN)	Pendiente del Terreno (P)	Densidad de Población (DP)	Proximidad a cuerpos de agua (PVF)
Proximidad a vías terrestres (PVT)	1	1/3	1/3	3	1
Intensidad de Luces Nocturnas (LN)	3	1	1	5	5
Pendiente del Terreno (P)	3	1	1	3	3
Densidad de Población (DP)	1/3	1/5	1/3	1	3
Proximidad a cuerpos de agua (PVF)	1	1/5	1/3	1/3	1
Suma por columna	8.3	2.7	3	12.3	13

Fuente: DANE. Censo Económico

- Con la suma anterior se calcula la matriz normalizada, dividiendo cada elemento de la MCP entre la suma de su columna respectiva. La figura 13 muestra el resultado para la MCP de ejemplo.

Figura 13. MCP normalizada

MCN	PVT	LN	P	DP	PVF
PVT	0,12	0,12	0,11	0,24	0,08
LN	0,36	0,37	0,33	0,41	0,38
P	0,36	0,37	0,33	0,24	0,23
DP	0,04	0,07	0,11	0,08	0,23
PVF	0,12	0,07	0,11	0,03	0,08

$$MCN_{ij} = \frac{MCP_{ij}}{\sum^m MCP_{i.}}$$

Fuente: DANE. Censo Económico

- Finalmente, se obtiene el vector de pesos con el promedio aritmético por fila de la matriz normalizada. Este vector genera la fracción o porcentaje de cada variable involucrada en la evaluación multicriterio. La figura 14 ilustra el resultado del vector de pesos para la MCP de ejemplo.

Figura 14. Vector de pesos con las ponderaciones o porcentajes

Vector promedio (W)
0,13
0,37
0,31
0,11
0,08

$$W_i = \frac{\sum_{l=1}^m MCN_{il}}{m}$$

Fuente: DANE. Censo Económico

2.3.6.7 Cálculo de la razón de consistencia RC

Para cada MCP obtenida se calcula la razón de consistencia con el fin de verificar que no haya contradicciones en las comparaciones pareadas. Para ello, en primer lugar, se calcula el producto matricial de la MCP por el vector de pesos, obteniendo un nuevo vector columna. A continuación, los elementos de este vector se dividen uno a uno entre los elementos del vector de pesos, como se ilustra en la figura 15.

Figura 15. Producto matricial y cociente de vectores

Producto matrices (MCP * W)	Producto de matrices /Vector Pesos (W_i)
0,76	5,67
2,02	5,47
1,65	5,37
0,57	5,35
0,43	5,24

Fuente: DANE. Censo Económico

A continuación, se calcula el promedio del vector cociente, lo que se conoce como valor propio o autovalor MAX y se calcula el índice de consistencia IC, dado por la ecuación:

$$IC = \frac{\lambda_{max} - n}{n - 1} \quad (1)$$

En la ecuación, n es el número de variables y MAX es el autovalor calculado previamente.

Finalmente, se calcula la razón de consistencia RC como el cociente entre el IC y un índice aleatorio IA, correspondiente al índice de consistencia promedio de 100.000 matrices aleatorias de la misma dimensión. Este valor se encuentra en tablas, y para un n = 5, IA = 1,115.

La razón de consistencia se evalúa de acuerdo con el número de variables. Para n = 5, RC debe ser menor de 0,1 para que se acepte que la matriz de comparación de pesos no presenta contradicciones. En caso contrario, se deben evaluar nuevamente las comparaciones entre variables.

$$RC = \frac{IC}{IA} \quad (2)$$

2.3.6.8 Asignación de los pesos a las variables y cálculo de la suma ponderada

Una vez completados los pasos anteriores y verificada la consistencia de las matrices, se asigna a cada variable su respectivo peso o ponderación.

Para el presente ejercicio, se recogieron las opiniones de doce (12) expertos (10 profesionales del DANE y dos externos), quienes diligenciaron

la MCP en dos escenarios: 1. Municipios con predominancia de vías terrestres y 2. Municipios con predominancia de vías fluviales (los cuales se encuentran en el Andén Pacífico y en la Amazonia). Los resultados de las matrices promediadas (medias geométricas), vectores de pesos y razones de consistencia para cada escenario se muestran en las figuras 16 y 17, respectivamente.

Figura 16. Media geométrica de MCP y vector de pesos para escenario 1

MCP	Proximidad a vías terrestres (PVT)	Intensidad de Luces Nocturnas (LN)	Pendiente del Terreno (P)	Densidad de Población (DP)	Proximidad a cuerpos de agua (PCA)	Vector pesos (W)
Proximidad a vías terrestres (PVT)	1	3.564	5.106	1.162	5.034	39.75%
Intensidad de Luces Nocturnas (LN)	0.281	1	2.011	0.449	1.390	12.72%
Pendiente del Terreno (P)	0.196	0.497	1	0.195	0.848	6.84%
Densidad de Población (DP)	0.860	2.229	5.141	1	3.860	31.82%
Proximidad a cuerpos de agua (PCA)	0.199	0.719	1.179	0.259	1	8.34%

$$RC_{\text{ESCENARIO 1}} = 0.0350 \text{ (Matriz consistente)}$$

Fuente: DANE. Censo Económico

Figura 17. Media geométrica de MCP y vector de pesos para escenario 2

MCP	Proximidad a vías terrestres (PVT)	Intensidad de Luces Nocturnas (LN)	Pendiente del Terreno (P)	Densidad de Población (DP)	Proximidad a cuerpos de agua (PCA)	Vector pesos (W)
Proximidad a vías terrestres (PVT)	1	0.785	1.052	0.291	0.179	7.92%
Intensidad de Luces Nocturnas (LN)	1.274	1	1.473	0.338	0.211	9.93%
Pendiente del Terreno (P)	0.950	0.679	1	0.197	0.166	6.91%
Densidad de Población (DP)	3.439	2.961	5.075	1	0.559	29.20%
Proximidad a cuerpos de agua (PCA)	5.576	4.729	6.039	1.790	1	46.03%

$$RC_{\text{ESCENARIO 2}} = 0.0033 \text{ (Matriz consistente)}$$

Fuente: DANE. Censo Económico

Ahora bien, en el andén Pacífico se observa que las vías terrestres tienen tanta importancia como las vías fluviales, por lo que fue necesario crear un tercer escenario en donde se calcule la media geométrica de las matrices resultantes de los escenarios 1 y 2. El resultado se muestra en la figura 18.

Figura 18. Media geométrica de MCP y vector de pesos para escenario 3

MCP	Proximidad a vías terrestres (PVT)	Intensidad de Luces Nocturnas (LN)	Pendiente del Terreno (P)	Densidad de Población (DP)	Proximidad a cuerpos de agua (PCA)	Vector pesos (W)
Proximidad a vías terrestres (PVT)	1	1.673	2.318	0.581	0.950	20.61%
Intensidad de Luces Nocturnas (LN)	0.598	1	1.721	0.389	0.542	13.02%
Pendiente del Terreno (P)	0.431	0.581	1	0.196	0.375	7.98%
Densidad de Población (DP)	1.720	2.569	5.108	1	1.469	35.64%
Proximidad a cuerpos de agua (PCA)	1.052	1.844	2.668	0.681	1	22.75%

$$RC_{\text{ESCENARIO 3}} = 0.0023 \text{ (Matriz consistente)}$$

Fuente: DANE. Censo Económico

Adicionalmente se consideraron otros dos escenarios: municipios sin cuerpos de agua y municipios sin vías. Los primeros se encuentran en su totalidad en la región Andina, y en este caso se elimina la variable Proximidad a cuerpos de agua. El segundo caso se encuentra en municipios del Andén Pacífico y la Amazonía, y se elimina la variable Proximidad a vías.

Los vectores finales de pesos se presentan en la tabla 12.

Tabla 12. Vectores de pesos por escenario

Variables	Porcentajes				
	Terrestre	Fluvial	Combinado	Sin vías	Sin cuerpos agua
Proximidad vías terrestres	39.75	7.92	20.61	0	42.91
Intensidad luces nocturnas	12.72	9.93	13.02	10.65	14.22
Pendiente del terreno	6.84	6.91	7.98	7.27	7.42
Densidad de población	32.36	29.20	35.64	31.97	35.45
Proximidad cuerpos de agua	8.34	46.03	22.75	50.11	0

Fuente: DANE. Censo Económico

Los porcentajes se asignan a la variable respectiva y se realiza la suma ponderada para obtener el ráster resultante.

2.3.6.9 Asignación de pesos a las variables y obtención de la suma ponderada

Antes de otorgar el porcentaje obtenido en la matriz de comparación a cada variable, es necesario establecer el puntaje para las clases al interior de cada una. Las clases en sí mismas se comportan como subcriterios, por lo cual también se les puede aplicar el método AHP. Al considerar que los valores de las clases disminuyen gradualmente de la clase de mayor valor a la de menor valor, se tiene la MCP y los pesos obtenidos que se muestran en la figura 19.

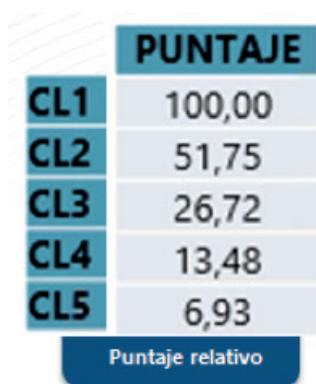
Figura 19. MCP de clases dentro de las variables y vector de pesos



Fuente: DANE. Censo Económico

No obstante, el vector de pesos debe normalizarse al ser un subcriterio, dividiendo por el mayor valor (0,5082 de la clase 1). Al hacer esto y multiplicar por 100, quedan expresados los pesos en términos de puntajes dentro de cada variable (Figura 20).

Figura 20. Vector de puntajes dentro de cada variable



Fuente: DANE. Censo Económico

Así, la clase a la que se le asigne mayor valor quedará con un puntaje de 100, la siguiente con un puntaje de 51,75 y así sucesivamente.

A continuación, se muestra el cálculo de la suma ponderada para el municipio de Fusagasugá (región Andina, predominio de vías terrestres tipo 1, 2, 3). En primer lugar, se muestran los rangos de las variables y sus puntajes internos.

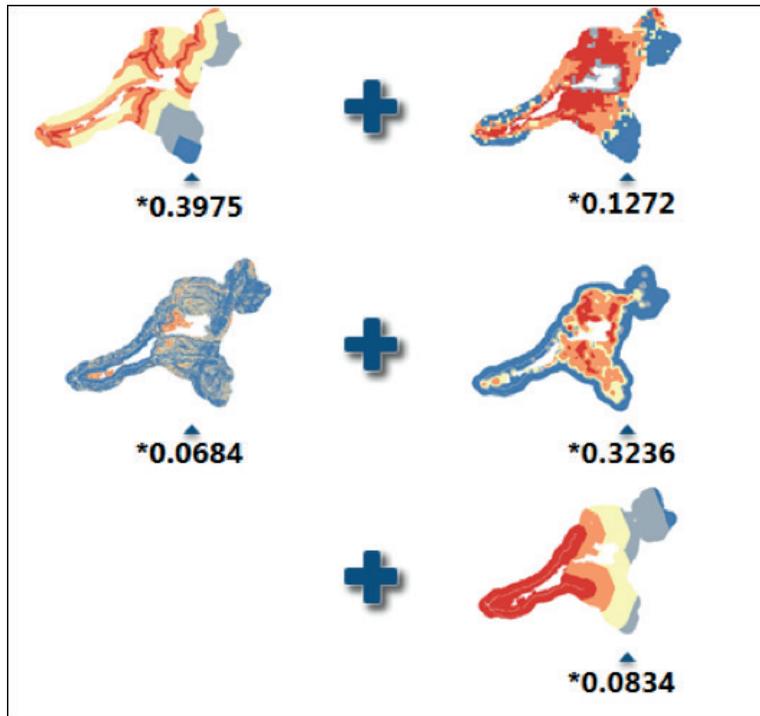
Figura 21. Rangos de las variables y puntajes en el municipio de Fusagasugá

Variable/Peso	Valores/Rangos originales	Calificación
Proximidad a vías terrestres 39.75%	0 – 159 metros	100.00
	159 – 955 metros	51.75
	955 – 2513 metros	26.72
	2513 – 5503 metros	13.48
	Mayor a 5503 metros	6.93
Luces Nocturnas 12.72%	0 – 0,414	6.93
	0,414 – 0,469	13,48
	0,469 – 0,883	26,72
	0,883 – 4,015	100.00
	4,015 – 27,720	51,75
Pendiente 6.84%	0 – 3%	100.00
	3 – 7%	51.75
	7 – 12%	26.72
	12 – 25%	13.48
	Mayor a 25%	6.93
Densidad de población 32.36%	0.00 – 2.40	6.93
	2.40 – 11.65	13.48
	11.65 – 47.29	26.72
	47.29 – 184.65	51.75
	184.65 – 714.07	100.00
Proximidad a cuerpos de agua 8.34%	0 – 1487 metros	100.00
	1487 – 3666 metros	51.75
	3666 – 6288 metros	26.72
	6288 – 9828 metros	13.48
	Mayor a 9828 metros	6.93

Fuente: DANE. Censo Económico

Enseguida se muestran las imágenes ráster que se emplean para la suma ponderada, multiplicando cada una por su correspondiente porcentaje.

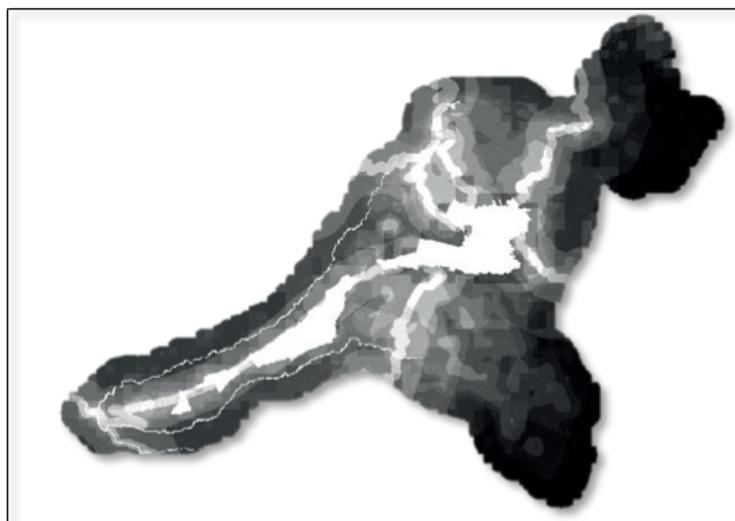
Figura 22. Ilustración de la suma ponderada de imágenes ráster en el municipio de Fusagasugá



Fuente: DANE. Censo Económico

El resultado de la suma ponderada se muestra en la figura 23.

Figura 23. Resultado del cálculo de la suma ponderada de las variables en el municipio de Fusagasugá



Fuente: DANE. Censo Económico

Para determinar el valor del umbral o valor de píxel mínimo a partir del cual se considera área óptima o potencial, se tomó una muestra de 32 municipios, con características físicas heterogéneas, reclasificando el ráster resultante en dos categorías mediante la técnica de intervalos iguales, el cual divide los valores en rangos de igual tamaño. La tabla 13 muestra los resultados del umbral para cada municipio y el porcentaje de este umbral con respecto al valor máximo de cada ráster.

Tabla 13. Resultados de umbrales con el método de intervalos iguales para diferentes municipios

Municipio	Nombre	Umbral_equal_interval	Maximo_ráster	Porcentaje_ei
44847	Uribia	53.463	100	53.46%
81794	Tame	48.803	90.682	53.82%
25279	Fómeque	54.152	100	54.15%
27001	Quibdó	53.462	100	53.46%
15693	Santa Rosa de Viterbo	50.409	100	50.41%
76834	Tuluá	49.857	92.788	53.73%
86749	Sibundoy	55.331	100	55.33%
66170	Dosquebradas	46.316	82.895	55.87%
08832	Tubará	55.262	100	55.26%
17001	Manizales	53.462	100	53.46%
70508	Ovejas	50.409	92.242	54.65%
19001	Popayán	50.668	93.865	53.98%
95001	San José del Guaviare	51.452	95.979	53.61%
85410	Tauramena	53.879	100	53.88%
50001	Villavicencio	54.765	100	54.77%
52240	Chachagüi	48.207	89.489	53.87%
68418	Los Santos	51.725	95.979	53.89%
13244	El Carmen de Bolívar	51.452	95.979	53.61%
15861	Ventaquemada	50.682	93.892	53.98%
18247	El Doncello	53.463	100	53.46%

20250	El Paso	53.687	100	53.69%
23162	Cereté	54.765	100	54.77%
25183	Chocontá	52.342	100	52.34%
25290	Fusagasugá	53.736	100	53.74%
41132	Campoalegre	50.409	93.892	53.69%
47980	Zona Bananera	54.765	100	54.77%
50124	Cabuyaro	51.383	92.788	55.38%
54498	Ocaña	53.463	100	53.46%
63690	Salento	53.463	100	53.46%
73124	Cajamarca	53.462	100	53.46%
91001	Leticia	51.452	95.979	53.61%
05318	Guarne	53.736	100	53.74%

Fuente: DANE. Censo Económico

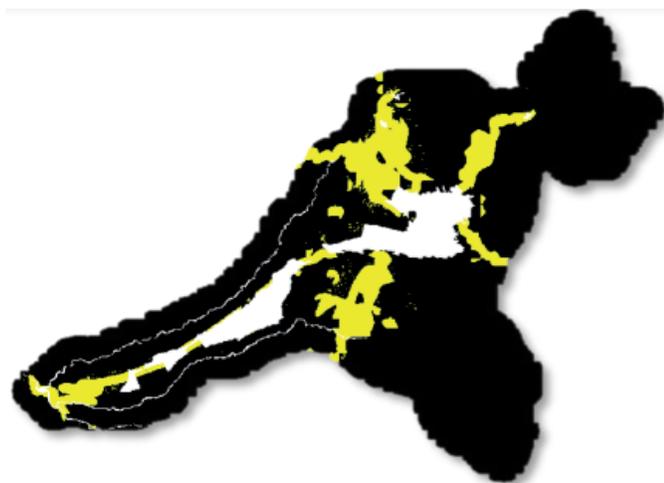
El promedio del porcentaje de los umbrales con respecto al máximo es de 53,90%. Aproximando el umbral al múltiplo de 10 superior más cercano, se tiene 60% del valor máximo del ráster como el límite a partir del cual se define un área óptima.

En conclusión, se considera el área óptima o área potencial de actividad económica en el rural disperso aquella que presenta valores de píxel iguales o superiores al 60% del valor máximo del ráster. Por otra parte, este valor límite asegura la participación de las variables más importantes en cada escenario.

El ráster resultante de Fusagasugá presenta un rango de valores desde 7 hasta 100. Esta imagen se reclasifica de manera que los valores de píxel iguales o superiores a 60 queden representados de un color distinto (amarillo) a los que sean inferiores a 60 (negro). El resultado se muestra en la figura 24.



Figura 24. Resultado de la reclasificación de la suma ponderada en el municipio de Fusagasugá



Fuente: DANE. Censo Económico

El área óptima de Fusagasugá se encuentra en las proximidades de las vías y en los lugares con mayor densidad poblacional. Finalmente, esta área óptima se extrae como un polígono y se cruza con las ACER obtenidas en el municipio, excluyendo aquellas que no se intersectan y tienen un número de establecimientos igual o inferior a cuatro, como se observa en la figura 25, donde las ACER excluidas aparecen en color azul claro.

Figura 25. Intersección de ACER con área óptima calculada mediante AHP



Fuente: DANE. Censo Económico

2.3.7 Validación y codificación de ACER

Los polígonos de referencia de las ACER seleccionados con el procedimiento de intersección con el área óptima de variables exógenas deben pasar por un proceso de validación y ajuste a las unidades del Marco Geoestadístico Nacional del DANE. El objetivo de este proceso es integrar los polígonos generados mediante los geoprocursos con los niveles definidos para el Marco Geoestadístico Nacional (MGN), a nivel de Sección Rural, mediante la asignación de un código y una delimitación precisa a cada ACER. El proceso completo de codificación se fundamenta en los siguientes argumentos:

- Permite la integración de información de otras operaciones estadísticas a nivel de sección rural, como es el caso de variables del CNPV.
- Facilita la georreferenciación de las unidades económicas en el siguiente nivel de desagregación geográfica (sección rural), para los casos en los cuales las condiciones atmosféricas o climáticas no alcancen el nivel de precisión del GPS requerido.
- Permite la georreferenciación de las unidades económicas a nivel de sección rural para los casos en que se trabaje en formulario en papel.

Las etapas de este macroproceso son: i) validación de polígonos de referencia con imágenes de satélite, ii) delimitación con respecto a límites prediales, iii) límites del Marco Maestro Rural Agropecuario (Conglomerados MMRA) y límites de sección rural; y, iv) asignación de codificación.

2.3.7.1 Validación de polígonos de referencia de ACER con imágenes de satélite

Es necesario validar la existencia de las áreas de concentración económica encontradas mediante los geoprocursos anteriores, empleando una fuente paralela como son las imágenes de satélite.

Pueden darse 4 situaciones en esta revisión:

- **El polígono de referencia del ACER se encuentra sobre un área de concentración económica.** Esto se evidencia por la presencia de infraestructura (techos de viviendas, bodegas, etc.) y la cercanía a una vía terrestre, fluvial o marítima. De igual forma, el polígono puede verificarse con otras fuentes como Google Maps®.
- **El polígono de referencia del ACER se encuentra desplazado por la georreferenciación de las fuentes de información.** En los polígonos de referencia, hay siempre un grado de incertidumbre dado por las fuentes de información de entrada. Así, puede darse el caso de que un polígono se encuentre desplazado con respecto a **la concentración verdadera, por lo que es necesario mover el polígono a la posición que le corresponde.**
- **El polígono de referencia del ACER se encuentra sobre un área despoblada o coincide con un centro poblado o cabecera municipal.** En cualquiera de estos tres casos, el polígono de referencia se elimina de la base de datos.
- **No hay imágenes disponibles o hay cobertura de nubes.** En este caso, no queda otra opción que asumir la validez del polígono de referencia.

2.3.7.2 Delimitación de ACER

La delimitación de los polígonos de referencia de ACER validados en el paso anterior requiere contar con varios niveles de información adicionales:

- Marco Operativo del CNPV 2018, con delimitación predial.
- Marco Maestro Rural Agropecuario 2019, con delimitación de conglomerados.
- Marco Geoestadístico Nacional (MGN), a

nivel de sección rural.

En primer lugar, se verifica si el polígono de ACER está dividido entre dos secciones rurales. De ser así, cada parte del ACER en secciones diferentes tendrá una codificación distinta.

A continuación, se construye el ACER como tal, empleando los niveles de información

mencionados y respetando, en lo posible, los límites prediales.

En caso de que se tengan predios demasiado extensos, la delimitación se realiza teniendo en cuenta límites naturales y bordes de construcciones, así como la ubicación de los establecimientos.

2.3.7.3 Codificación de ACER

Una vez se ha delimitado con precisión el área de concentración económica rural, se le asigna un código de dieciséis (16) dígitos: 14 posiciones correspondientes a los niveles geográficos del MGN utilizado para las operaciones estadísticas DANE y 2 posiciones al código asignado para las áreas de concentración económica rural.

En la figura 26 se observa un ejemplo de codificación en el municipio de Fusagasugá (Cundinamarca).

Figura 26. Ejemplo de código para área de concentración económica rural (ACER)



Fuente: DANE. Censo Económico

2.3.8 Indicadores de importancia relativa de las ACER

Una vez identificadas, delimitadas y estructuradas las ACER, es necesario establecer una serie de indicadores que permitan determinar su importancia relativa al interior de cada municipio, con miras a priorizar las de mayor jerarquía durante los operativos de campo y en futuras investigaciones donde se incluyan otras variables a partir de los establecimientos (número de empleados, consumo de servicios públicos, etc.) que contribuyan a mejorar la caracterización de la actividad económica en el área rural.

En los siguientes incisos se define el concepto de indicador y se presentan algunas propuestas de indicadores para las ACER.

2.3.8.1 Definición de indicador

De acuerdo con el Departamento Nacional de Planeación, un indicador es una “expresión cuantitativa observable y verificable, que permite describir características, comportamientos o fenómenos de la realidad. Esto se logra a través de la medición de una variable o una relación de variables”²⁵. Técnicamente, “es una función de una o más variables que conjuntamente miden una característica o un atributo de los individuos en estudio”²⁶.

Las principales características de un indicador son las siguientes²⁷:

Simplificación de la realidad.

Un indicador integra y resume diferentes dimensiones de un tema, creando una imagen de contexto que facilita la interpretación del fenómeno estudiado. Esta se considera la principal característica de un indicador: resumir en un valor numerosos aspectos interrelacionados.

Capacidad de medición.

Implica la comparación de los valores obtenidos en el indicador con respecto a un patrón o escala establecido previamente.

Capacidad de comunicación.

Todo indicador debe transmitir información acerca de un tema en particular para facilitar la toma de decisiones.

La construcción de un indicador requiere dos condiciones fundamentales:

- Definición clara del atributo o característica que se desea medir, lo cual provee un sustento conceptual al indicador.
- Disponibilidad de información confiable para poder realizar la medición, lo que otorga validez al indicador.

Teniendo en cuenta lo anterior, se proponen cinco (5) indicadores para describir a las áreas de concentración económica rural (ACER) identificadas, validadas y estructuradas, con el fin de avanzar en el proceso de caracterización de la actividad económica en el rural disperso y difundir estos resultados en herramientas como geovisores o tablas.



²⁵ Dirección de Seguimiento y Evaluación de Políticas Públicas (2018). Guía para la construcción y análisis de indicadores. Departamento Nacional de Planeación DNP. Recuperado de https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Guia_para_elaborar_Indicadores.pdf

²⁶ Schuschny, A., Soto, H. (2009). Guía metodológica. Diseño de indicadores compuestos de desarrollo sostenible. Documento de Proyecto. Comisión Económica para América Latina y el Caribe CEPAL, GTZ. Recuperado de https://www.cepal.org/ilpes/noticias/paginas/9/35989/disenio_indicadores_compuestos_ddss.pdf

²⁷ Opcit.

Los indicadores propuestos son los siguientes:

2.3.8.2 Distancia relativa del ACER a un área urbana (DR)

2.3.8.2.1 Descripción general

Este indicador cuantifica la distancia estandarizada del centroide²⁸ del ACER al área urbana más cercana (cabecera municipal o centro poblado) medida en metros.

Se define distancia estandarizada como el cociente entre la distancia obtenida de la ACER al área urbana más cercana y la máxima distancia medida entre dos puntos extremos del municipio. Esta medida tiene como propósito caracterizar las ACER más lejanas y cercanas a un área urbana (no ubicada necesariamente en el municipio, por ejemplo, una cabecera municipal vecina).

2.3.8.2.2 Objetivo del indicador

El indicador DR tiene como objetivo una caracterización inicial de las ACER relativa a su distancia a la cabecera municipal o centros poblados, la cual ha mostrado ser un factor determinante tanto en el número de establecimientos económicos como en la extensión de la ACER en los municipios visitados durante las pruebas de campo.

2.3.8.2.3 Metodología de cálculo

El indicador DR se calcula como sigue:

$$DR = \frac{DC}{D_{max}} \times 100\%$$

Donde:

DC: distancia del centroide del ACER al área urbana más cercana (cabecera municipal o centro poblado) medida en metros

Dmax: distancia máxima entre dos puntos en el municipio, medida en metros.

Las distancias se calculan con una herramienta de sistemas de información geográfica y corresponden a distancias euclidianas (medidas en línea recta).

2.3.8.2.4 Unidad de medida del indicador

El indicador DR es un número adimensional expresado como porcentaje, ya que es el resultado del cociente de dos distancias medidas en metros.

2.3.8.2.5 Interpretación del valor del indicador

El indicador DR puede tomar valores de cero a cien (0,100).

Un valor cercano a cero indica que la ACER se encuentra más próxima a un área urbana.



²⁸ Se define el centroide o baricentro de un polígono como el punto tal que cualquier recta que pase por él divide al polígono en dos partes iguales (igual momento con respecto a la recta).

2.3.8.3 Participación de establecimientos económicos del ACER sobre el total en las ACER del municipio (P1)

2.3.8.3.1 Unidad de medida del indicador

Este indicador mide la participación o porcentaje del número de establecimientos económicos que pertenecen a un ACER determinada con respecto al total de establecimientos contenidos en las ACER del municipio.

Esta medida tiene como propósito ordenar las ACER de mayor a menor en función del número de establecimientos económicos de cada una con respecto al total dentro de todas las ACER.

2.3.8.3.2 Objetivo del indicador

El indicador P1 tiene como objetivo otorgar una participación o porcentaje a cada ACER en función del número de establecimientos económicos, permitiendo de esta manera organizar las ACER de mayor a menor participación, lo cual permite visualizar cuáles ACER tienen mayor porcentaje y, por tanto, mayor importancia relativa dentro del municipio.

2.3.8.3.3 Metodología de cálculo

El indicador P1 se calcula como sigue:

$$P1 = \frac{TE_{ACER}}{TE_{ACER\ MPIO}} \times 100$$

Donde.

TE_{ACER} : total de establecimientos de la ACER evaluada

$TE_{ACER\ MPIO}$: suma de los establecimientos en todas las ACER del municipio

2.3.8.4 Unidad de medida del indicador

El indicador P1 es un número adimensional expresado en porcentaje, ya que es el resultado del cociente de dos conteos de establecimientos económicos.

2.3.8.3.5 Interpretación del valor del indicador

El indicador P1 puede tomar valores de cero a cien (0,100).

Si el valor de P1 es mayor, significa que la participación del número de establecimientos de la ACER es más grande con respecto al total.

2.3.8.4 Participación de establecimientos económicos del ACER sobre el total general del rural disperso del municipio (P2)

2.3.8.4.1 Descripción general

Este indicador mide la participación o porcentaje del número de establecimientos económicos que pertenecen a un ACER determinada con respecto al total de establecimientos económicos en el rural disperso (clase 3) del municipio.

Esta medida tiene como propósito ordenar las ACER de mayor a menor en función del número de establecimientos económicos de cada una con respecto al total del rural disperso.

En este indicador, los establecimientos económicos del área rural no delimitada como el ACER también participan en el cálculo.

2.3.8.4.2 Objetivo del indicador

El indicador P2 tiene como objetivo otorgar una participación o porcentaje a cada ACER en función del número total de establecimientos económicos en el rural disperso, permitiendo de esta manera organizar las ACER de mayor a menor participación y teniendo en cuenta los establecimientos que

no quedaron incluidos dentro de las ACER. De esta forma, es posible comparar la participación de una ACER determinada con el porcentaje de establecimientos económicos que no se encuentran en ninguna ACER.

2.3.8.4.3 Metodología de cálculo

El indicador P2 se calcula como sigue:

$$P2 = \frac{TE_{ACER}}{TE_{MPIO CL 3}} \times 100$$

Donde:

TE_{ACER} : total de establecimientos de la ACER evaluada

$TE_{MPIO CL 3}$: total de establecimientos en la clase 3 (rural disperso) del municipio

2.3.8.4.4 Unidad de medida del indicador

El indicador P2 es un número adimensional expresado en porcentaje, ya que es el resultado del cociente de dos conteos de establecimientos económicos.

2.3.8.4.5 Interpretación del valor del indicador

El indicador P2 puede tomar valores de cero a cien (0,100).

Si el valor de P2 es mayor, significa que la participación del número de establecimientos de la ACER es más grande con respecto al total general.

2.3.8.5 Distancia-Peso de la ACER en el municipio (DW1)

2.3.8.5.1 Descripción general

Con este indicador se pretende combinar la información provista por el indicador DR y por la

participación de cada ACER en el total del número de establecimientos en las ACER, mediante una razón (cociente) de los indicadores P1 y DR.

Esta medida tiene como propósito la caracterización y jerarquización de las ACER a través una medida que alcanza valores máximos cuando la ACER se encuentra cerca de un área urbana y presenta una alta participación del número de establecimientos con respecto al total de establecimientos contenidos en las ACER del municipio.

2.3.8.5.2 Objetivo del indicador

Con este indicador se pretende combinar la información provista por el indicador DR y por la participación de cada ACER en el total del número de establecimientos en las ACER, mediante una razón (cociente) de los indicadores P1 y DR.

Esta medida tiene como propósito la caracterización y jerarquización de las ACER a través una medida que alcanza valores máximos cuando la ACER se encuentra cerca de un área urbana y presenta una alta participación del número de establecimientos con respecto al total de establecimientos contenidos en las ACER del municipio.

2.3.8.5.3 Metodología de cálculo

El indicador DW1 se calcula como sigue:

$$DW1 = \frac{P1}{DR}$$

Donde:

DR: indicador de distancia relativa

P1: participación de establecimientos económicos de la ACER relativa al total de EE.EE en las ACER del municipio.

2.3.8.5.4 Unidad de medida del indicador

El indicador DW1 es un número adimensional.

2.3.8.5.5 Interpretación del valor del indicador

El indicador DW1 puede tomar valores positivos.

Se interpreta de la siguiente forma:

- Valor cercano a cero: la ACER se encuentra muy lejos de cualquier área urbana y tiene pocos establecimientos.
- Valor positivo muy grande: la ACER se encuentra cerca de un área urbana.
- Para ACER ubicadas a la misma distancia de un área urbana, un valor mayor de una con respecto a la otra se refleja en un mayor número de establecimientos en la primera.

2.3.8.6 Distancia-Peso de la ACER en el municipio relativo al rural disperso (DW2)

2.3.8.6.1 Descripción general

Con este indicador se pretende combinar la información provista por el indicador DR y por la participación de cada ACER en el total del número de establecimientos en la clase 3, mediante una razón (cociente) de los indicadores P2 y DR.

Estamediatienecomopropósitolacaracterización y jerarquización de las ACER a través una medida que alcanza valores máximos cuando la ACER se encuentra cerca de un área urbana y contiene gran cantidad de establecimientos económicos, que reflejan una alta participación relativa al total de establecimientos en el rural disperso (clase 3).

2.3.8.6.2 Objetivo del indicador

El indicador DW2 tiene como objetivo la

jerarquización de las ACER, considerando dos variables: participación de la ACER con respecto al total de establecimientos económicos del rural disperso (clase 3) del municipio y su distancia relativa al área urbana más cercana.

El cociente entre estas dos variables proporciona un indicador compuesto, cuyo valor es creciente conforme disminuye la distancia a un área urbana y aumenta el número de establecimientos económicos en el ACER, permitiendo así organizar las ACER de mayor a menor importancia dentro del municipio con respecto a la facilidad de llegar a ella desde la cabecera municipal (distancia relativa) y su participación en el total de establecimientos del municipio.

2.3.8.6.3 Metodología de cálculo

El indicador DW2 se calcula como sigue:

$$DW2 = \frac{P2}{DR}$$

Donde:

DR: indicador de distancia relativa

P2: participación de establecimientos económicos de la ACER relativa al total de EE.EE en las ACER del municipio.

2.3.8.6.4 Unidad de medida del indicador

El indicador DW2 es un número adimensional.

2.3.8.6.5 Interpretación del valor del indicador

El indicador DW2 puede tomar valores positivos.

Se interpreta de la siguiente forma:

- Valor cercano a cero: la ACER se encuentra muy lejos de cualquier área urbana y tiene una participación de establecimientos muy baja con respecto al total del municipio.

- Entre mayor sea el valor, la ACER se encuentra más cerca de un área urbana.
- Para ACER ubicadas a la misma distancia de un área urbana, un valor mayor de una con respecto a la otra se refleja en un mayor número de establecimientos en la primera.

3. RESULTADOS



3.1. Resultados del municipio de Ubaté (Cundinamarca)

Para el municipio de Ubaté (código DIVIPOLA 25843), se tienen las siguientes estadísticas generales con respecto a los establecimientos económicos en el área rural.

Tabla 14. Parámetros del municipio de Ubaté (Cundinamarca)

Parámetro	Valor	
Número de establecimientos	CNPV 2018	70
	MAPAS RUTA COLOMBIA	25
	EST. EDUCATIVOS	11
	IPS	4
	BOOKING	4
	OPEN STREET MAP	3
	TOTAL	122
Promedio Vecino Más Cercano*	357,26 m	
Mediana Vecino Más Cercano*	207,77 m	
Distancia máxima*	1.799,12 m	
Distancia mínima*	11,09 m	
Desviación estándar*	375,24 m	
Número mínimo de establecimientos en el ACER	3	

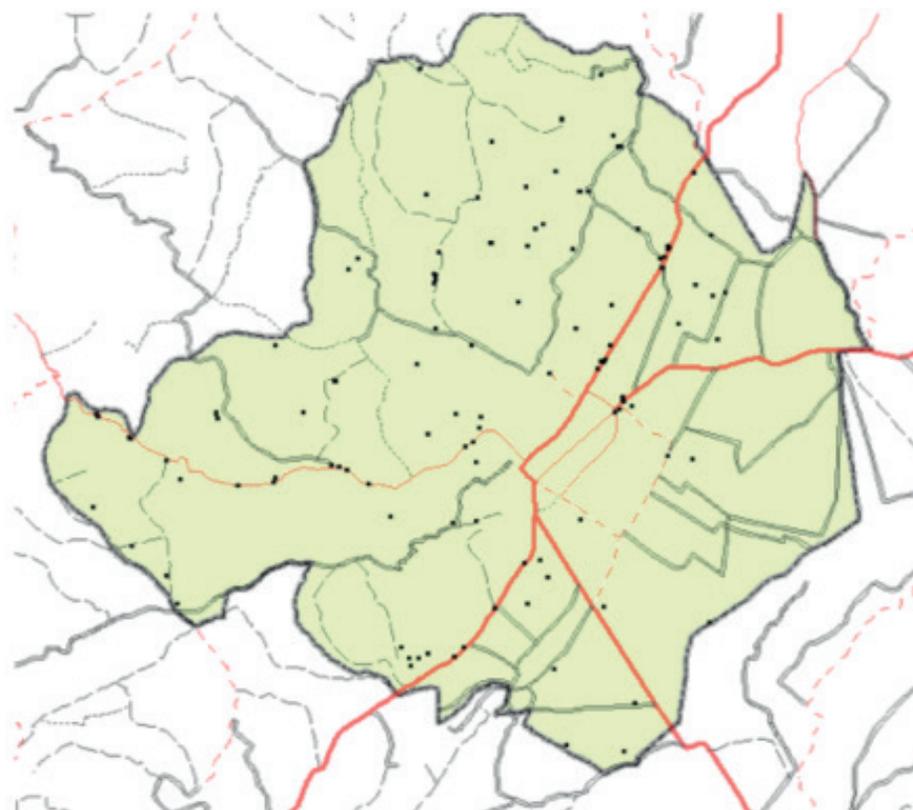
*Referida a establecimientos económicos

Fuente: DANE. Censo Económico

El valor de la desviación estándar (375,24 m) de las distancias al establecimiento vecino más cercano es muy similar al promedio, lo que unido al diagrama de caja y bigotes de distancias confirma que la mediana es la magnitud de tendencia central más representativa, al existir valores atípicos cercanos a cero y también mayores a un kilómetro.

A continuación, se muestra la ubicación de los establecimientos económicos en el municipio de Ubaté.

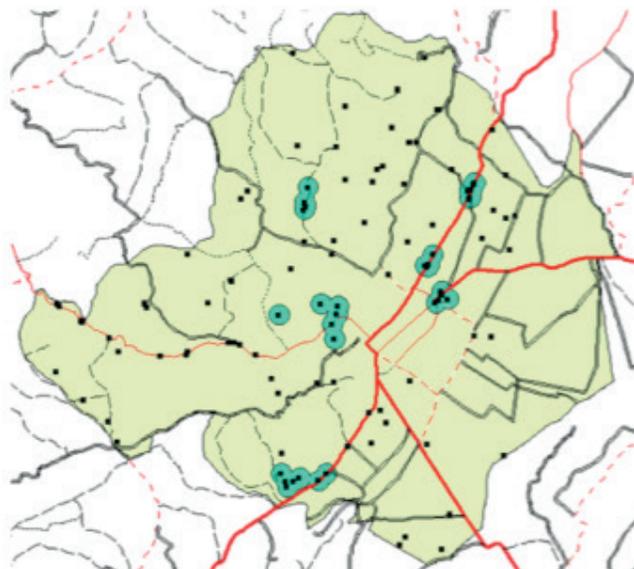
Figura 27. Ubicación de establecimientos económicos en clase 3 (rural disperso) en el municipio de Ubaté (Cundinamarca)



Fuente: DANE. Censo Económico

Al ejecutar el geoproceto de área de influencia (Buffer) sobre los establecimientos económicos, seguido de la disolución de los polígonos y la selección de aquellos que contengan más de cinco establecimientos, se obtiene la capa mostrada en la figura 28 en color azul.

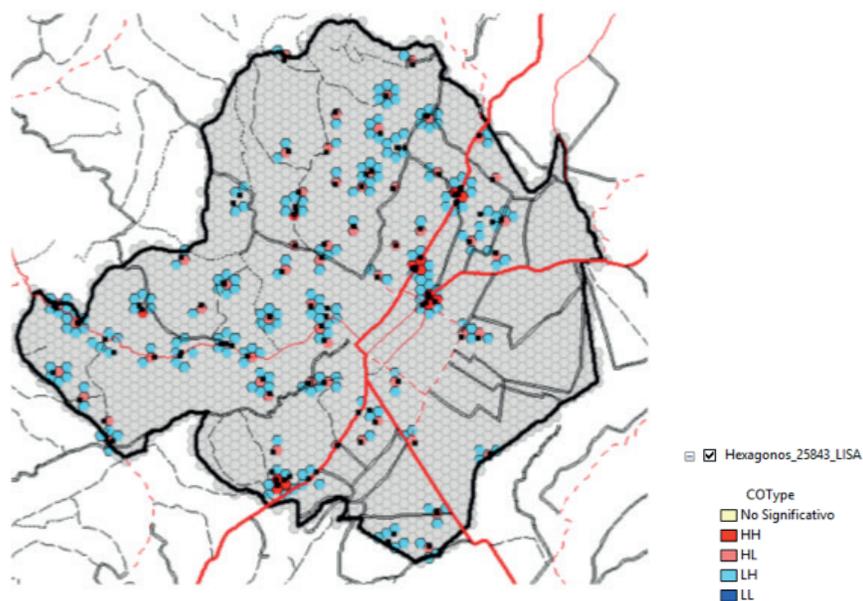
Figura 28. Ubicación de áreas de influencia o buffer en el municipio de Ubaté (Cundinamarca)



Fuente: DANE. Censo Económico

Al correr el geoproceto de índices locales de autocorrelación espacial (LISA) se obtienen los clusters (concentraciones) y outliers (valores atípicos) mostrados en la figura 29.

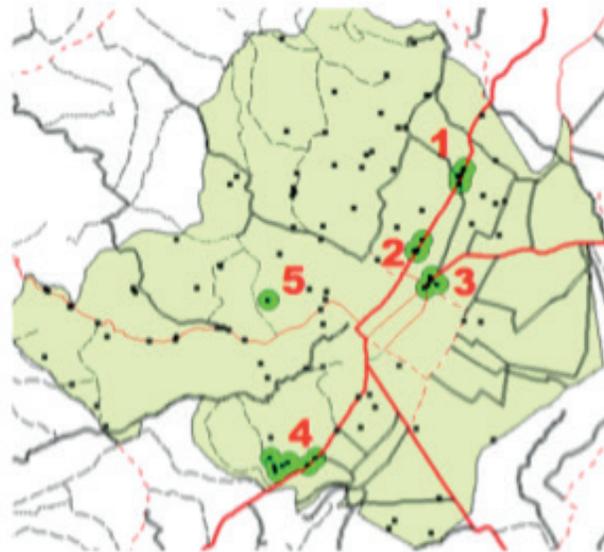
Figura 29. Ubicación de clusters y outliers de establecimientos económicos en Ubaté (Cundinamarca)



Fuente: DANE. Censo Económico

Finalmente, se determinan cuáles áreas de influencia se intersectan con clusters HH o outliers HL, obteniendo la capa de polígonos de referencia de las ACER.

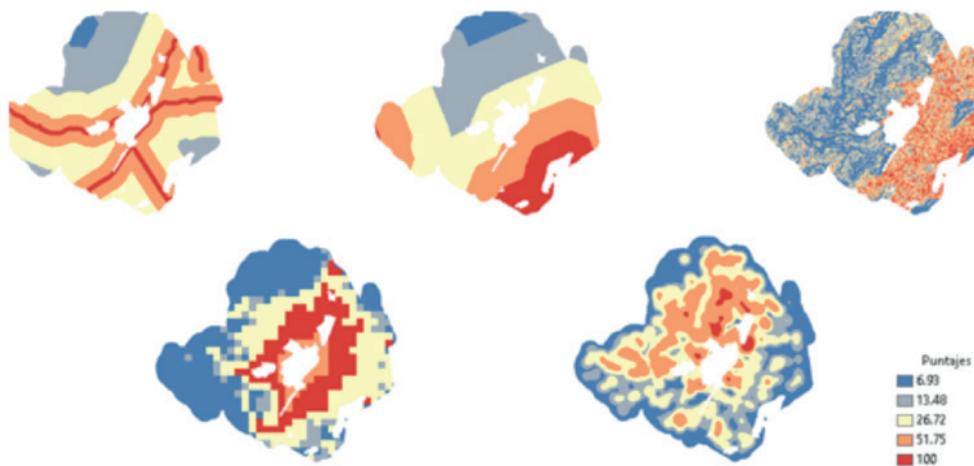
Figura 30. Ubicación de polígonos de referencia de ACER en Ubaté (Cundinamarca)



Fuente: DANE. Censo Económico

La validación de ACER en Ubaté se realiza con el área óptima de las variables exógenas y las imágenes de satélite. Para Ubaté aplican los rangos de la región Andina, con predominancia de vías 1, 2 y 3.

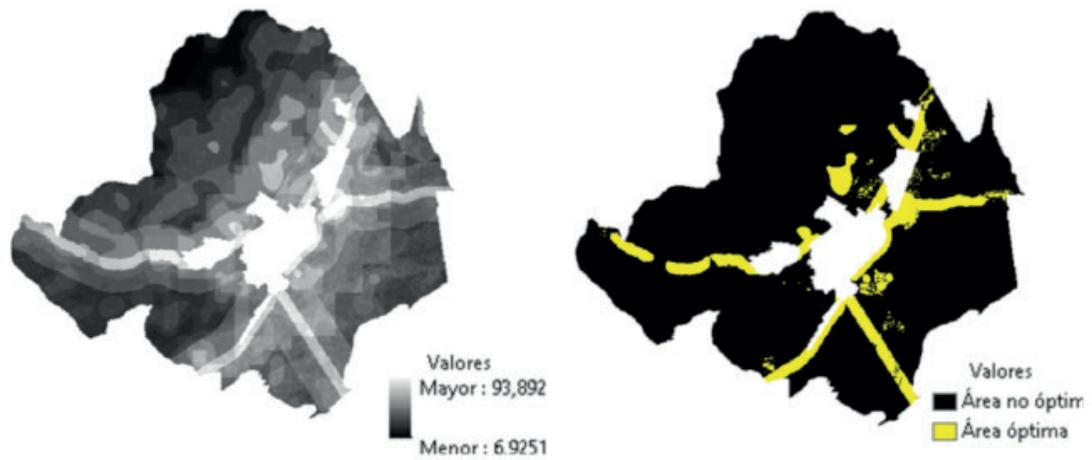
Figura 31. Imágenes ráster para calcular el área óptima o potencial de actividad económica rural en Ubaté.



Fuente: DANE. Censo Económico

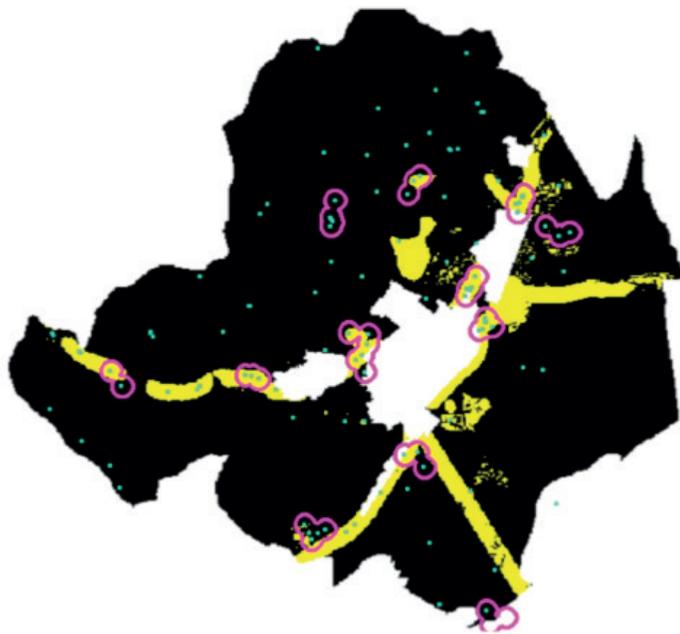
De izquierda a derecha: proximidad a vías terrestres, luces nocturnas, proximidad a cuerpos de agua, densidad de población y pendiente del terreno.

Figura 32. Imágenes ráster de resultado de la ponderación y selección del área óptima de actividad económica en el rural disperso



Fuente: DANE. Censo Económico

Figura 33. Área óptima intersectada con ACER del municipio de Ubaté



Fuente: DANE. Censo Económico

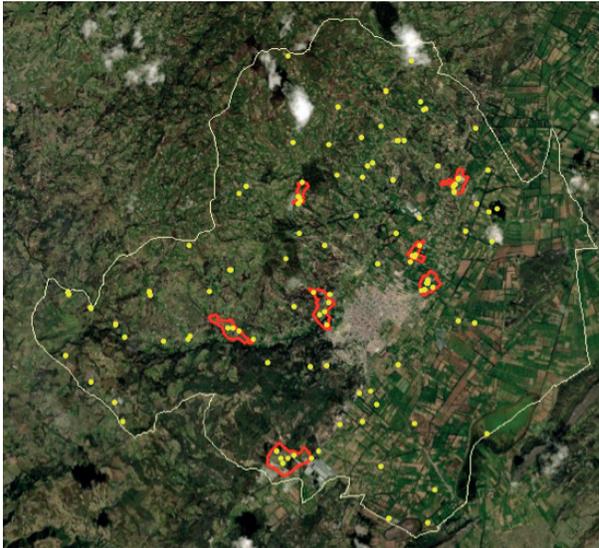
Tabla 15. Algunos resultados de validación de ACER en el municipio de Ubaté (Cundinamarca)

ID ACER	Observaciones	Imagen
1	Se observa evidencia de presencia y concentración de establecimientos económicos (techos de viviendas, infraestructura).	
2	Se observa evidencia de presencia y concentración de establecimientos económicos (techos de viviendas, infraestructura).	
3	Se observa evidencia de presencia y concentración de establecimientos económicos (techos de viviendas, infraestructura).	
4	Se observa una alta dispersión de los establecimientos económicos, y algunos de ellos no se encuentran cerca de construcciones o infraestructura.	
5	No se observa presencia de infraestructura o viviendas en esta área ni en sitios cercanos, por lo que se concluye que esta área no corresponde realmente a un ACER.	

Fuente: DANE. Censo Económico

El resultado de la estructuración se observa en la Figura 34.

Figura 34. Resultado de estructuración de las ACER del municipio de Ubaté (Cundinamarca)



COD MPIO	ACER	Geocodigo
25843	01	250025843399999999999948955501
25843	02	250025843399999999999948955502
25843	03	250025843399999999999948955703
25843	04	250025843399999999999948955704
25843	05	250025843399999999999948955705
25843	06	250025843399999999999948955606
25843	07	250025843399999999999948955807

Fuente: DANE. Censo Económico

Al realizar el cálculo de los indicadores, se obtienen los siguientes resultados que se presentan en la Tabla 15.

Tabla 16. Resultados de indicadores de importancia de ACER para el municipio de Ubaté (Cundinamarca)

COD_MPIO	ACER	NUM_EST	DIST_CP_CERCANO	DR	P1	P2	DW1	DW2
25843	1	3	526,3	3,76%	8,57%	2,46%	2,28	0,65
25843	2	5	134,2	0,96%	14,29%	4,10%	14,90	4,27
25843	3	5	210,8	1,51%	14,29%	4,10%	9,49	2,72
25843	4	5	121,3	0,87%	14,29%	4,10%	16,49	4,73
25843	5	7	329,7	2,36%	20,00%	5,74%	8,49	2,44
25843	6	5	1831,2	13,08%	14,29%	4,10%	1,09	0,31
25843	7	5	982,5	7,02%	14,29%	4,10%	2,04	0,58

Fuente: DANE. Censo Económico

- **Distancia máxima al interior del municipio:** 13,980 metros.
- **Total de establecimientos en las ACER:** 35.
- **Total de establecimientos en el municipio:** 122.

Los resultados de los indicadores permiten establecer que, en términos de la distancia a áreas urbanas y número de establecimientos, la ACER de mayor puntaje en los indicadores compuestos DW1 y DW2 es la 04, la cual se encuentra a la menor distancia de un área urbana (121,3 metros).

Por otra parte, los menores puntajes corresponden a la ACER 06, la cual se encuentra a la mayor distancia de un área urbana (1.831,2 metros).

4. CONCLUSIONES



La metodología desarrollada en el presente documento permite la identificación y delimitación de las áreas de concentración económica rural, o ACER, áreas consideradas representativas de la actividad económica no agropecuaria en el rural disperso o clase 3.

La metodología tiene en cuenta las particularidades del área rural de cada municipio, al tomar como medida de tendencia central a la mediana de la distancia de los establecimientos a su vecino más cercano. Esta elección también otorga robustez a los cálculos posteriores.

Las áreas obtenidas mediante los algoritmos de área de influencia (buffer) e índices locales de autocorrelación espacial (LISA) se validan mediante la intersección con áreas potenciales de actividad económica en el rural disperso generadas a partir de un análisis multicriterio jerárquico (AHP) y posteriormente se delimitan en el proceso de estructuración, ajustándolas tanto a divisiones del MGN como a límites naturales, y asignándoles una codificación consistente con el MGN.

Las ACER obtenidas se jerarquizan a través de indicadores compuestos, teniendo en cuenta tanto su

número de establecimientos como la distancia a las áreas urbanas más cercanas (cabeceras municipales, centros poblados).

Lo anterior permite establecer un marco censal rural consistente tanto con las fuentes secundarias de información como con el Marco Geoestadístico Nacional. Sin embargo, debe indicarse que las ACER pueden o no abarcar los establecimientos rurales de inclusión forzosa, tales como las estaciones de servicio, las zonas francas o los parques industriales.

El trabajo desarrollado para la conformación de las ACER es la base para próximas investigaciones que mejoren la delimitación existente y puedan emplearse en otras operaciones estadísticas, teniendo en cuenta aquellos municipios donde los establecimientos económicos en el área rural constituyen un alto porcentaje del total de establecimientos, donde la actividad económica rural es de mayor importancia o más representativa que la del área urbana, o en donde se observa la existencia de corredores turísticos que explican en gran parte la dinámica económica tanto del municipio como de sus vecinos.

www.dane.gov.co
