



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

**Centro Nacional de Investigación
y Desarrollo Tecnológico**

Tesis de Doctorado

**Desarrollo de un modelo matemático con enfoque en
la Toma de Decisiones Multicriterio:
Caso de estudio accesibilidad a servicios**

presentada por

M. en I. Odette Alejandra Pliego Martínez

como requisito para la obtención del grado de
Doctora en Ciencias de la Computación

Directora de tesis

Dra. Alicia Martínez Rebollar

Codirector de tesis

Dr. Hugo Estrada Esquivel

Cuernavaca, Morelos, México. Agosto de 2025.

	ACEPTACIÓN DE IMPRESIÓN DEL DOCUMENTO DE TESIS DOCTORAL		Código: CENIDET-AC-006-D20
	Referencia a la Norma ISO 9001:2008 7.1, 7.2.1, 7.5.1, 7.6, 8.1, 8.2.4		Revisión: 0
			Página 1 de 1

Cuernavaca, Mor., a 26 de mayo de 2025

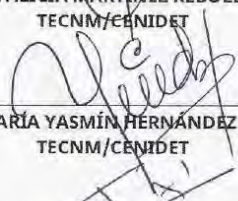
DR. CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
P R E S E N T E

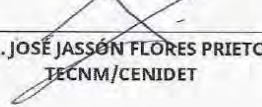
AT'n: DR. JUAN GABRIEL GONZÁLEZ SERNA
PRESIDENTE DEL CLAUSTRO DOCTORAL DEL
DEPARTAMENTO DE CIENCIAS COMPUTACIONALES

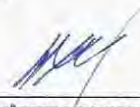
Los abajo firmantes, miembros del Comité Tutorial de la estudiante **M.C. ODETTE ALEJANDRA PLIEGO MARTÍNEZ** manifiestan que después de haber revisado el documento de tesis titulado **"Desarrollo de un Modelo Matemático con Enfoque en la Toma de Decisiones Multicriterio: Caso de Estudio Accesibilidad a Servicios"**, realizado bajo la dirección de la **Dra. Alicia Martínez Rebollar** y codirección del **Dr. Hugo Estrada Esquivel**, el trabajo se **ACEPTA** para proceder a su impresión.

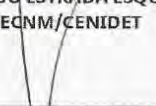
ATENTAMENTE
Excelencia en Educación Tecnológica®
"Conocimiento y Tecnología al Servicio de México"


DRA. ALICIA MARTÍNEZ REBOLLAR
TECNM/CENIDET


DRA. MARÍA YASMÍN HERNÁNDEZ PÉREZ
TECNM/CENIDET


DR. JOSÉ JASSÓN FLORES PRIETO
TECNM/CENIDET


DR. HUGO ESTRADA ESQUIVEL
TECNM/CENIDET


DR. JAVIER ORTIZ HERNÁNDEZ
TECNM/CENIDET


DR. ELÍAS NEFTALÍ ESCOBAR GÓMEZ
TECNM/ IT. TUXTLA GUTIÉRREZ

c.c.p: C Verónica Sotelo Boyas/ Jefa del Departamento de Servicios Escolares
c.c.p: C Née Alejandro Castro Sánchez / Jefe del Departamento de Ciencias Computacionales
c.c.p: Expediente

CENIDET-AC-006-D20

Rev. 0



Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO



Centro Nacional de Investigación y Desarrollo Tecnológico
Subdirección Académica

Cuernavaca Mor, 04/junio/2025

Oficio No. SAC/122/2025

Asunto: Autorización de impresión de tesis

ODETTE ALEJANDRA PLIEGO MARTÍNEZ
CANDIDATA AL GRADO DE DOCTORA
EN CIENCIAS DE LA COMPUTACIÓN
P R E S E N T E

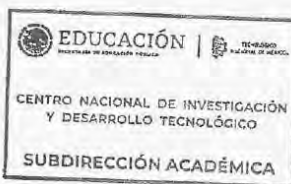
Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado **"Desarrollo de un Modelo Matemático con Enfoque en la Toma de Decisiones Multicriterio: Caso de Estudio Accesibilidad a Servicios"**, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE

Excelencia en Educación Tecnológica®
"Conocimiento y Tecnología al Servicio de México"

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO



c.c.p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/lmz



2025
Año de
La Mujer
Indígena

Interior Internado Palmira S/N, Col. Palmira,
C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 4104,
e-mail: acad_cenidet@tecnm.mx tecnm.mx | cenidet.tecnm.mx



Dedicatoria

Dedico esta tesis a Dios, cuya presencia ha iluminado mi camino siempre, brindándome sabiduría, fortaleza y esperanza en cada paso.

A mi esposo, Ernesto de la Cruz, cuya pasión por el conocimiento y motivación constante me inspiró y me acompañaron a lo largo de este proceso. Juntos recorrimos el camino del estudio, enfrentando desafíos y superando obstáculos. Su apoyo incondicional, su orientación durante las etapas más complicadas fueron fundamentales para que pudiera avanzar. A él le debo no solo gran parte de este logro, sino también el apoyo para seguir aprendiendo y creciendo.

A mis padres, Gemma Martínez y Santiago Pliego, por su respaldo y apoyo me guiaron para ir por el camino del conocimiento.

Agradecimientos

Al Maestro en Tecnologías de la Información Ernesto de la Cruz, por su apoyo incondicional, su motivación inquebrantable, su dedicación constante y su perseverancia para superar cada desafío que se presentó, contribuyendo de manera fundamental al logro de esta meta.

A mi directora de tesis, la Dra. Alicia Martínez, por compartir generosamente su conocimiento, por su valiosa orientación y por el gran apoyo brindado tanto en el ámbito académico como personal. Su acompañamiento constante fue fundamental para la realización de este trabajo.

A mi codirector el Dr. Hugo Estrada, por su atención, guía y conocimiento, que complementó el trabajo de mi directora para sacar adelante este proyecto. También le agradezco por su apoyo en los aspectos personales durante este proceso.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por la beca otorgada, la cual fue fundamental para la realización de estos estudios.

Al Tecnológico Nacional de México (TecNM) por facilitar sus instalaciones, a través del Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), lo cual fue fundamental durante el desarrollo de este trabajo.

A mi revisor externo, Dr. Elías Escobar, por sus valiosas sugerencias y observaciones, así como por el tiempo dedicado, los cuales contribuyeron significativamente a la mejora de esta investigación.

A los doctores Nimrod González, Eddie Clemente y Jonathan Villanueva por permitirme tomar sus asignaturas como oyente y por compartir generosamente su conocimiento; lo cual contribuyó de manera significativa a enriquecer este trabajo.

A mi tía Tayde Martínez por su interés y por preguntar cómo avanzaba este proceso, brindándome ánimo en cada etapa.

Al Actuario Aarón Martínez Rangel, un ser querido que siempre mostró un profundo interés por la ciencia; cuya curiosidad lo impulsó por ir más allá y dejar una huella imborrable. Aunque ya no está físicamente, es seguro que, de haber estado presente, habría seguido cada paso de este trabajo con entusiasmo, interés y orgullo.

A mis padres, por su orientación y quienes me señalaron el camino a seguir en el mundo del conocimiento.

Al señor Cliserio de la Cruz y María Nicolás por su apoyo en diversos aspectos, por su interés genuino y por estar siempre al pendiente en cada etapa de este proceso.

A Refugia de la Cruz y Dalia Ramírez por estar siempre al pendiente, por su apoyo constante y su motivación para salir adelante.

Agradezco a mi comité tutorial por sus valiosos comentarios y observaciones, los cuales contribuyeron significativamente a enriquecer este trabajo.

A todas aquellas personas dentro del CENIDET, que de alguna brindaron su apoyo para seguir adelante tanto académicamente como en la parte administrativa.

Resumen

En los procesos de planeación y gestión, la selección de la mejor alternativa frente a problemas complejos exige considerar múltiples variables, frecuentemente interrelacionadas y, en ocasiones, conflictivas. La Toma de Decisiones Multicriterio ofrece un marco metodológico eficaz para describir el comportamiento del problema y facilitar la evaluación y selección de alternativas en contextos de conflicto. No obstante, una de las etapas críticas de esta metodología es la asignación de pesos a las variables, la cual suele depender del juicio experto. Este enfoque, aunque valioso, puede introducir limitaciones importantes, como la subjetividad, el sesgo cognitivo y la dificultad para conformar perfiles expertos adecuados.

Como respuesta a estas limitaciones, una de las principales contribuciones de esta investigación es la construcción de un método de ponderación objetiva, denominado PCA-Entropía, para la asignación de pesos a las variables, basado en la información contenida en los datos. Este método combina la Teoría del Análisis de Componentes Principales con la Teoría de la Entropía, permitiendo asignar pesos que integran tanto la varianza explicada como el nivel de información independiente de cada variable; lo que refuerza su robustez. Además, el enfoque propuesto es menos sensible a los efectos de la normalización o estandarización. Como resultado, se obtiene una herramienta sólida, coherente y reproducible, adecuada para abordar problemas multicriterio complejos y de naturaleza heterogénea.

Otra contribución relevante de esta investigación es el desarrollo de un modelo matemático multicriterio, aplicado a un caso de estudio centrado en la accesibilidad a servicios dentro de una zona geográfica específica. El modelo fue construido mediante técnicas de Estadística, Minería de datos, Toma de Decisiones Multicriterio y Sistemas de Información Geográfica. Como parte fundamental de su metodología de solución, se implementó el método de ponderación PCA-Entropía para la asignación objetiva de pesos a las variables. El modelo resultante, denominado Indicador de Accesibilidad a Servicios, integra siete variables agrupadas en tres criterios o dimensiones clave: vivienda, educación y espacial. La combinación de estas variables permitió una evaluación integral y cuantitativa del nivel de accesibilidad territorial, facilitando la selección de zonas con distintos grados de acceso a servicios.

Los resultados obtenidos en la etapa de validación mostraron coherencia y concordancia con modelos existentes, lo cual fue corroborado mediante coeficientes de correlación. Para evaluar su confiabilidad y estabilidad, se llevaron a cabo análisis de sensibilidad y ablación. Ambos análisis confirmaron que el modelo mantiene un comportamiento consistente ante distintos escenarios y que las tres dimensiones propuestas aportan un enfoque integral y complementario, indispensable para una evaluación precisa y holística del acceso a servicios.

Palabras clave: Modelo matemático multicriterio, Indicador de Accesibilidad a Servicios, Accesibilidad a servicios, Método de ponderación de variables PCA-Entropía, Asignación de pesos a las variables, Toma de Decisiones Multicriterio, Sistemas de Información Geográfica, Ciudad de México, Selección de variables relevantes.

Abstract

In planning and management processes, selecting the best alternative when facing complex problems requires considering multiple variables, often interrelated and, in some cases, conflicting. Multicriteria Decision Making provides an effective methodological framework to describe the behavior of such problems and to support the evaluation and selection of alternatives in conflict-laden contexts. However, one of the critical stages in this methodology is the assignment of weights to variables, which is typically based on expert judgment. While this approach can be valuable, it also introduces significant limitations, such as subjectivity, cognitive bias, and difficulties in assembling suitably qualified expert panels.

In response to these limitations, one of the main contributions of this research is the development of an objective weighting method, called PCA-Entropy, for assigning weights to variables based on the information contained in the data. This method combines Principal Component Analysis with Entropy Theory, allowing the assignment of weights that integrate both the explained variance and the independent information level of each variable, thereby enhancing the method's robustness. In addition, the proposed approach is less sensitive to the effects of data normalization or standardization. As a result, it offers a solid, coherent, and reproducible tool suitable for addressing complex and heterogeneous multicriteria problems.

Another important contribution of this research is the development of a mathematical multicriteria model, applied to a case study focused on accessibility to services within a specific geographic area. The model was built using techniques from Statistics, Data Mining, Multicriteria Decision Making, and Geographic Information Systems. As a core element of its methodological solution, the PCA-Entropy weighting method was implemented to assign objective weights to the variables. The resulting model, named the Service Accessibility Index, integrates seven variables grouped into three key criteria or dimensions: housing, education, and spatial. The combination of these variables enabled a comprehensive and quantitative evaluation of territorial accessibility, supporting the identification of areas with varying degrees of access to services.

The validation results demonstrated coherence and agreement with existing models, as confirmed by correlation coefficients. To assess the model's reliability and stability, sensitivity and ablation analyses were conducted. Both analyses confirmed that the model exhibits consistent behavior under different scenarios, and that the three proposed dimensions provide an integrated and complementary perspective, essential for a precise and holistic evaluation of service accessibility.

Keywords: Multicriteria mathematical model, Service Accessibility Index, Accessibility to services, PCA-Entropy variable weighting method, Variable weight assignment, Multicriteria Decision Making, Geographic Information Systems, Mexico City, Relevant variable selection.

Contenido

Dedicatoria	iii
Agradecimientos	iv
Resumen	v
Abstract.....	vi
Lista de figuras	ix
Lista de tablas	xi
Capítulo 1. Introducción.....	1
1.1. Panorama general.....	2
1.2. Motivación	3
1.3. Planteamiento del problema	3
1.4. Objetivos de la investigación.....	5
1.4.1. Objetivo general	5
1.4.2. Objetivos específicos.....	5
1.5. Hipótesis de investigación.....	5
1.6. Metodología de solución de la investigación	6
1.7. Contribuciones a las Ciencias Computacionales.....	6
1.8. Organización de la tesis.....	7
Capítulo 2. Estado del arte.....	8
2.1. Introducción	9
2.2. Antecedentes de los modelos de accesibilidad a servicios.....	9
2.3. Modelos y métodos de accesibilidad a servicios.....	13
2.4. Modelos de accesibilidad a servicios mediante la Toma de Decisiones Multicriterio	28
2.5. Análisis del estado del arte.....	34
Capítulo 3. Marco teórico	35
3.1. Introducción	36
3.2. Toma de Decisiones Multicriterio	36
3.2.1. Toma de Decisiones Multicriterio.....	38
3.2.2. Elementos de la Toma de Decisiones Multicriterio.....	39
3.3. Técnicas para la asignación de pesos en modelos multicriterio.....	44
3.3.1. Análisis de Componentes Principales	44
3.3.2. Método de Entropía	46
3.4. Técnicas de identificación de variables relevantes	47
3.4.1. Factor de Inflación de la Varianza.....	47
3.4.2. Correlación lineal	48

3.4.3. Análisis de Componentes Principales.....	49
3.4.4. Análisis Factorial.....	49
3.5. Sistemas de información geográfica.....	49
3.6. Accesibilidad a servicios.....	50
3.6.1. Definición de accesibilidad a servicios.....	51
3.6.2. Componentes de la accesibilidad.....	51
3.6.3. Definición de accesibilidad para esta investigación	52
Capítulo 4. Desarrollo del modelo matemático multicriterio.....	54
4.1. Introducción.....	55
4.2. Construcción de un método de ponderación de variables con enfoque objetivo para la asignación de pesos.....	55
4.2.1. Asignación de pesos de las variables.....	55
4.2.2. Desarrollo del Método de ponderación de variables PCA-Entropía	57
4.2.3. Caso de aplicación del método de ponderación de variables PCA-Entropía	64
4.3. Metodología de solución para el desarrollo del modelo matemático multicriterio	72
4.3.1. Caso de estudio utilizado	76
4.4. Desarrollo del modelo matemático multicriterio	78
4.4.1. Preprocesamiento de los datos.....	79
4.4.2. Construcción del modelo multicriterio.....	105
4.5. Conclusión.....	130
Capítulo 5. Validación e interpretación del modelo matemático multicriterio	132
5.1. Introducción	133
5.1.1. Análisis comparativo	133
5.1.2. Análisis de sensibilidad	140
5.1.3. Análisis de ablación	153
5.2. Interpretación del modelo.....	155
5.2.1. Conjunto de datos evaluado	158
5.2.2. Representación gráfica	160
5.3. Conclusión	172
Capítulo 6. Conclusiones y trabajos futuros	173
6.1. Conclusiones generales.....	174
6.1.1. Trabajos futuros.....	175
6.1.2. Logros obtenidos.....	176
Referencias	179

Lista de figuras

Figura 3.1. Metodología de la Toma de Decisiones Multicriterio.....	39
Figura 3.2. Relación entre metas, objetivos, criterios y variables.....	40
Figura 3.3. Elementos de los problemas de Toma de Decisiones Multicriterio.....	41
Figura 3.4. Relación entre los componentes de la accesibilidad.....	53
Figura 4.1. Metodología de solución de la investigación.....	55
Figura 4.2. Proceso de asignación dentro de la Toma de Decisiones Multicriterio.....	57
Figura 4.3. Método ponderación de variables PCA-Entropía.....	59
Figura 4.4. Pesos obtenidos mediante el método PCA-Entropía.....	69
Figura 4.5. Esquema de la metodología de solución.....	76
Figura 4.6. Elementos para el caso de estudio.....	81
Figura 4.7. Distribución geográfica de la Ciudad de México.....	83
Figura 4.8. Ubicación de las escuelas de nivel primaria.....	84
Figura 4.9. Cálculo de la ruta en la plataforma de Geoapify.....	85
Figura 4.10. Cálculo de las variables de distancia y tiempo.....	86
Figura 4.11. Ejemplo 1 de la Interpolación de vecindad espacial.....	88
Figura 4.12. Ejemplo 2 de la Interpolación de vecindad espacial.....	89
Figura 4.13. Histogramas de frecuencia de las variables. Parte 1.....	91
Figura 4.14. Histogramas de frecuencia de las variables. Parte 2.....	92
Figura 4.15. Histogramas de frecuencia de las variables. Parte 3.....	92
Figura 4.16. Diagramas de caja de las variables. Parte 1.....	93
Figura 4.17. Diagramas de caja de las variables. Parte 2.....	93
Figura 4.18. Diagramas de caja de las variables. Parte 3.....	94
Figura 4.19. Mapa de coropletas de las variables Parte 1.....	94
Figura 4.20. Mapa de coropletas de las variables. Parte 2.....	95
Figura 4.21. Mapa de coropletas de las variables. Parte 3.....	95
Figura 4.22. Mapa de coropletas de las variables. Parte 4.....	96
Figura 4.23. Mapa de coropletas de las variables. Parte 5.....	96
Figura 4.24. Mapa de coropletas de las variables. Parte 6.....	97
Figura 4.25. Mapa de coropletas de las variables. Parte 7.....	97
Figura 4.26. Resultados de la prueba de densidad. Parte 1.....	100
Figura 4.27. Resultados de la prueba de densidad. Parte 2.....	100
Figura 4.28. Resultados de la prueba de densidad. Parte 3.....	101
Figura 4.29. Resultados de la prueba de normalidad. Parte 1.....	101
Figura 4.30. Resultados de la prueba de normalidad. Parte 2.....	102
Figura 4.31. Resultados de la prueba de normalidad. Parte 3.....	102
Figura 4.32. Matriz de diagramas de dispersión de la correlación.....	108
Figura 4.33. Mapa de calor de la correlación.....	109
Figura 4.34. Porcentaje de varianza explicada.....	111
Figura 4.35. Gráfico de las Importancia relativa de las variables.....	112
Figura 4.36. Resultados de la prueba de Kaiser-Meyer-Olkin.....	113
Figura 4.37. Número óptimo de factores.....	114
Figura 4.38. Resultados del Análisis Factorial.....	114
Figura 4.39. Resultados del Análisis Factorial de manera gráfica.....	115
Figura 4.40. Variables relevantes clasificados por dimensión.....	116
Figura 4.41. Método ponderación de variables PCA-Entropía.....	117
Figura 4.42. Árbol de decisión del modelo multicriterio.....	130
Figura 4.43. Pesos relativos de los criterios en un ranking.....	130

Figura 5.1. Modelo multicriterio de accesibilidad a servicios por localidad propuesto	137
Figura 5.2. Indicador de oportunidades acumuladas por localidad	138
Figura 5.3. Indicador basado en la distancia por localidad.....	138
Figura 5.4. Índice de marginación por localidad	138
Figura 5.5. Modelo multicriterio de accesibilidad a servicios por municipio propuesto	138
Figura 5.6. Indicador de acumuladas por municipio	139
Figura 5.7. Indicador basado en la distancia por municipio	139
Figura 5.8. Índice de marginación por localidad	139
Figura 5.9. Comparación del modelo de accesibilidad propuesto y el índice de marginación.....	139
Figura 5.10. Histogramas SBASC: datos reales y simulados.....	142
Figura 5.11. Histogramas de OVHAC: datos reales y simulados.....	142
Figura 5.12. Histogramas de PGRAPROES: datos reales y simulados	142
Figura 5.13. Histogramas de PVPH_SNBIEN: datos reales y simulados	143
Figura 5.14. Histogramas de PVPH_SINTIC: datos reales y simulados.....	143
Figura 5.15. Histogramas de NUM_ESC: datos reales y simulados.....	143
Figura 5.16. Histogramas de SBASC: datos reales y simulados.....	144
Figura 5.17. Distribución por categorías del análisis con técnicas no paramétricas ...	145
Figura 5.18. Histogramas de las variables con distribución uniforme. Parte 1.....	146
Figura 5.19. Histogramas de las variables con distribución uniforme. Parte 2.....	147
Figura 5.20. Histogramas de las variables con distribución uniforme. Parte 3.....	147
Figura 5.21. Distribución por categorías del análisis con simulación estadística.....	148
Figura 5.22. Grafo con la relación de las variables	150
Figura 5.23. Simulación del escenario 1.....	152
Figura 5.24. Simulación del escenario 2	152
Figura 5.25. Simulación del escenario 3.....	153
Figura 5.26. Histograma de los intervalos de clase clasificador mediante Jenks.....	156
Figura 5.27. Árbol de decisión del indicador de accesibilidad a servicios	158
Figura 5.28. Nivel de accesibilidad en la Ciudad de México por localidad.....	162
Figura 5.29. Nivel de accesibilidad por localidad con las escuelas	163
Figura 5.30. Indicador de accesibilidad: municipio de Azcapotzalco	164
Figura 5.31. Indicador de accesibilidad: municipio de Coyoacán	164
Figura 5.32. Indicador de accesibilidad: del municipio de Cuajimalpa.....	165
Figura 5.33. Indicador de accesibilidad: municipio de Gustavo A. Madero	166
Figura 5.34. Indicador de accesibilidad: municipio de Iztacalco.....	166
Figura 5.35. Indicador de accesibilidad: municipio de Iztapalapa	167
Figura 5.36. Indicador de accesibilidad: municipio de La Magdalena Contreras.....	167
Figura 5.37. Indicador de accesibilidad: municipio de Milpa Alta	168
Figura 5.38. Indicador de accesibilidad: municipio de Álvaro Obregón.....	168
Figura 5.39. Indicador de accesibilidad por localidad: municipio de Tláhuac.....	169
Figura 5.40. Indicador de accesibilidad: municipio de Tlalpan.....	169
Figura 5.41. Indicador de accesibilidad: municipio de Xochimilco	170
Figura 5.42. Indicador de accesibilidad: municipio de Benito Juárez	170
Figura 5.43. Indicador de accesibilidad: municipio de Cuauhtémoc.....	171
Figura 5.44. Indicador de accesibilidad: municipio de Miguel Hidalgo	171
Figura 5.45. Indicador de accesibilidad: municipio de Venustiano Carranza.....	172

Lista de tablas

Tabla 2.1. Resumen de los antecedentes de la accesibilidad	12
Tabla 2.2. Resumen de los modelos y métodos de accesibilidad, parte 1	26
Tabla 2.3. Resumen de los modelos y métodos de accesibilidad, parte 2	27
Tabla 2.4. Resumen de los modelos y métodos de accesibilidad, parte 3	28
Tabla 2.5. Resumen modelos de accesibilidad con enfoque Multicriterio	33
Tabla 4.1. Asignación de pesos de las variables del índice de marginación	65
Tabla 4.2. Resumen estadístico del conjunto de datos	65
Tabla 4.3. Subconjunto de datos de la matriz de decisión	66
Tabla 4.4. Conjunto de datos transformado	66
Tabla 4.5. Matriz de covarianza	67
Tabla 4.6. Vector de valores propios	67
Tabla 4.7. Matriz de vectores propios	67
Tabla 4.8. Matriz de componentes principales	68
Tabla 4.9. Matriz de componentes principales normalizada	68
Tabla 4.10. Cálculo de la Entropía de cada componente	69
Tabla 4.11. Subconjunto de datos clasificado	70
Tabla 4.12. Resultado del coeficiente de correlación de Spearman	72
Tabla 4.13. Resultado de la matriz de confusión	72
Tabla 4.14. Descripción de las fuentes consultadas y la información extraída	81
Tabla 4.15. Número de escuelas de nivel primaria por municipio	83
Tabla 4.16. Ejemplos del número de escuelas por localidad	85
Tabla 4.17. Instancias de conjunto de datos para analizar	86
Tabla 4.18. Descripción de la integración del conjunto de datos	87
Tabla 4.19. Descripción del conjunto de datos, parte 1	90
Tabla 4.20. Descripción del conjunto de datos, parte 2	90
Tabla 4.21. Resultados de la prueba Shapiro-Wilks de algunas variables. Parte 1	99
Tabla 4.22. Resultados de la prueba Shapiro-Wilks de algunas variables. Parte 2	99
Tabla 4.23. Descripción del conjunto de datos transformado, parte 1	104
Tabla 4.24. Descripción del conjunto de datos transformado, parte 2	104
Tabla 4.25. Subconjunto de datos transformados	105
Tabla 4.26. Matriz del Factor de Inflación de la Varianza de cada criterio	107
Tabla 4.27. Variables seleccionadas a partir del factor de inflación de la varianza	107
Tabla 4.28. Matriz de correlación	108
Tabla 4.29. Prueba de significancia de algunas variables. Parte 1	110
Tabla 4.30. Prueba de significancia de algunas variables. Parte 2	110
Tabla 4.31. Variables seleccionadas a partir de la correlación lineal	110
Tabla 4.32. Importancia relativa de los variables por componente	111
Tabla 4.33. Variables seleccionadas de Análisis de Componentes Principales	112
Tabla 4.34. Variables seleccionadas a partir del Análisis Factorial	115
Tabla 4.35. Variables relevantes de mayor impacto	116
Tabla 4.36. Matriz de covarianza de la iteración 1	119
Tabla 4.37. Vector de valores propios de la iteración 1	119
Tabla 4.38. Matriz de vectores propios de la iteración 1	120
Tabla 4.39. Matriz de componentes principales de la iteración 1	120
Tabla 4.40. Matriz de componentes principales normalizada de la iteración 1	121
Tabla 4.41. Cálculo de la Entropía de cada componente de la iteración 1	121
Tabla 4.42. Pesos de los variables obtenidos en la iteración 1	121

Tabla 4.43. Subconjunto de datos evaluados y clasificados en la iteración 1.....	123
Tabla 4.44. Subconjunto de datos de prueba evaluados y clasificados en la iteración 1	124
Tabla 4.45. Métrica de la evaluación del rendimiento de la iteración 1	124
Tabla 4.46. Pesos obtenidos en cada iteración y pesos promedio.....	125
Tabla 4.47. Modelos multicriterio obtenidos en cada iteración	125
Tabla 4.48. Pesos obtenidos en cada iteración y pesos promedio	125
Tabla 4.49. Unidades de medida de las variables.....	127
Tabla 4.50. Unidades y valores de las variables de referencia	128
Tabla 5.1. Descripción de los modelos de referencia.....	135
Tabla 5.2. Resultados del análisis de concordancia y consistencia del modelo propuesto	136
Tabla 5.3. Descripción estadística de las variables originales.....	144
Tabla 5.4. Descripción estadística de las variables simuladas.....	144
Tabla 5.5 Resultados de las métricas del análisis de sensibilidad de técnicas no paramétricas.....	145
Tabla 5.6 Descripción estadística de las variables con distribución uniforme.....	147
Tabla 5.7 Resultados de las métricas análisis de sensibilidad en simulación estadística	148
Tabla 5.8. Matriz de correlación ponderada	150
Tabla 5.9. Medidas de centralidad del grafo	150
Tabla 5.10 Resultados del estudio de ablación	155
Tabla 5.11. Reglas de la escala numérica y cualitativa del nivel de accesibilidad	156
Tabla 5.12. Valor de decisión	157
Tabla 5.13. Conjunto de datos del modelo de accesibilidad.....	159
Tabla 5.14. Reporte del número de localidades por categoría	159
Tabla 5.15. Descripción del conjunto de datos de accesibilidad a las escuelas.....	160

Capítulo 1. Introducción

1.1. Panorama general

Este capítulo ofrece una introducción a la investigación, que se centra en el desarrollo de un modelo matemático multicriterio para evaluar la accesibilidad a los servicios. El modelo se basa en la Toma de Decisiones Multicriterio, utilizando un enfoque objetivo sustentado en los datos.

En la vida cotidiana y en la industria, tomar decisiones es una tarea constante, tanto para atender necesidades personales como para gestionar diversos procesos. El objetivo principal de un problema de decisión es seleccionar la mejor opción entre varias alternativas disponibles. En este sentido, los problemas de decisión bajo conflicto surgen cuando intervienen múltiples variables o factores que, además de ser diferentes entre sí, entran en conflicto o competencia simultánea; lo que complica la elección de una o más alternativas.

Los problemas de decisión bajo conflicto surgen cuando las decisiones no dependen de un único factor, sino de múltiples factores que pueden ser incompatibles o difíciles de equilibrar. En este contexto, la Toma de Decisiones Multicriterio se enfoca en resolver problemas con conflictos entre diferentes variables; permitiendo una evaluación equilibrada que integra tanto las variables que influyen en el problema como su importancia relativa.

Este análisis puede realizarse mediante un enfoque objetivo, sustentado en técnicas que se apoyan en la información cuantificable proporcionada por los datos. Por otro lado, también puede llevarse a cabo desde un enfoque subjetivo, basado en métodos que dependen del juicio o la opinión de expertos. Sin embargo, el enfoque objetivo facilita la implementación de técnicas matemáticas que aprovechan la información proporcionada por los datos para determinar el nivel de importancia de cada variable involucrada; mientras que el enfoque subjetivo puede verse limitado por la percepción individual y la posible falta de consenso entre expertos. Por lo tanto, el enfoque objetivo no solo facilita la evaluación de las alternativas, sino que también genera resultados más precisos y representativos de la realidad, contribuyendo a un proceso de toma de decisiones más equilibrado, riguroso y fundamentado.

La presente investigación aborda la accesibilidad a los servicios como caso de estudio y se enfoca en el desarrollo de un modelo matemático multicriterio; con el objetivo de describir la relación entre diversas variables y el comportamiento de la accesibilidad en una determinada zona geográfica. Asimismo, busca evaluar dicho comportamiento para identificar patrones y tendencias territoriales. El modelo propuesto pretende servir como una base sólida para la formulación de estrategias y políticas orientadas a una distribución más equitativa de los recursos en el futuro; tales como la asignación de nueva infraestructura educativa, programas de apoyo en zonas con mayores niveles de rezago, entre otros. Los resultados obtenidos pueden ser útiles para los ciudadanos, al facilitar decisiones informadas en su vida cotidiana relacionadas con el acceso a servicios; como para futuras investigaciones centradas en la accesibilidad y la equidad territorial.

1.2. Motivación

En la vida cotidiana, las personas deben satisfacer diversas necesidades esenciales, como el acceso al trabajo, la salud, la alimentación y la educación, entre otras. Estas necesidades influyen directamente en su calidad de vida y bienestar (INEGI, 1994), (González et al., 1989). Por tal motivo, la (ONU, 2025) y la Constitución Política de los Estados Unidos Mexicanos (2025) reconocen la alimentación, la salud y la educación como derechos fundamentales. Estos derechos son esenciales para el desarrollo personal y la generación de nuevas oportunidades para el futuro. Dichas necesidades, comúnmente denominadas servicios, son esenciales para garantizar una vida digna y fomentar tanto el desarrollo como el bienestar de las comunidades (Aguirre, 2017). Un acceso adecuado a estos servicios no solo mejora las condiciones individuales, sino que también fortalece el desarrollo humano y resalta la importancia de tomar decisiones fundamentadas para garantizar el bienestar integral de la población.

Sin embargo, el acceso a estos servicios no siempre es equitativo; ya que diversas variables como las sociales, demográficas, económicas y espaciales que influyen en su disponibilidad y generan disparidades que afectan de manera diferenciada a la población, limitando así su desarrollo y bienestar. Entre estas variables se encuentran la distancia a los servicios, el nivel de analfabetismo, la disponibilidad de bienes en la vivienda, entre otros factores. Analizar la distribución de los servicios en función de estas variables permite identificar áreas desatendidas y detectar brechas que impactan a distintos grupos poblacionales. Esta comprensión es fundamental para diseñar e implementar estrategias que mejoren la cobertura y garanticen un acceso oportuno y equitativo.

Con base en lo anterior, el análisis de accesibilidad se presenta como un problema complejo que debe ser abordado considerando una variedad de variables. La diversidad de factores involucrados requiere un enfoque integral que permita evaluar de manera equilibrada las distintas variables que afectan el acceso a los servicios.

1.3. Planteamiento del problema

Los problemas abordados por la Teoría de Decisiones, especialmente en contextos de conflicto, se caracterizan por la necesidad de seleccionar alternativas u opciones considerando múltiples variables que, con frecuencia, están interrelacionadas y presentan características contradictorias entre sí (Prawda, 1996), (Starr, 1971).

La Toma de Decisiones Multicriterio ofrece un marco metodológico que facilita la descripción del comportamiento del problema, así como la evaluación y selección de alternativas, mediante la consideración simultánea de múltiples variables; frecuentemente interrelacionadas y conflictivas entre sí. Esto permite enfrentar de manera sistemática los retos propios del proceso de decisión.

El proceso de Toma de Decisiones Multicriterio se estructura a través de varias etapas esenciales: (1) Definición y contextualización del problema, (2) Identificación de criterios (variables) y alternativas, (3) Asignación de pesos a las variables, (4) Aplicación de un modelo de agregación, (5) Evaluación de alternativas y análisis de resultados, (6) Análisis de sensibilidad y robustez, y (7) Implementación y seguimiento (Keshavarz-

Ghorabae et al., 2021). De estas etapas, la asignación de pesos representa un componente crítico; ya que determina la influencia o importancia relativa de cada variable sobre los resultados del modelo. Una asignación adecuada influye directamente en la evaluación y selección de alternativas.

Tradicionalmente, la asignación de pesos se ha realizado mediante métodos basados en el juicio de expertos, como el Proceso Analítico Jerárquico (AHP), el Proceso de Red Analítica (ANP) y el método Delphi. Estos enfoques permiten incorporar conocimiento especializado, al considerar la experiencia y el criterio de los expertos como una fuente clave para determinar la relevancia relativa de las variables involucradas (Michalus et al., 2015), (Rodríguez Cruz, 2020), (Azmoodeh et al., 2021), (García Valdés & Suárez Marín, 2013) y (Silvestre Ortiz, 2019), (Malczewski, 2017).

La aplicación de estos métodos implica una serie de etapas, entre las que se incluyen: (a) la selección de especialistas, (b) el diseño y aplicación de instrumentos para la recolección de información, (c) el procesamiento y análisis de los datos obtenidos, y (d) la determinación de los pesos correspondientes a cada variable, en función de los resultados y del nivel de consenso alcanzado.

Aunque estos enfoques han sido ampliamente utilizados, presentan limitaciones relevantes asociadas a la subjetividad, el sesgo cognitivo y la dependencia de la disponibilidad de expertos. Entre estas limitaciones se encuentran: la dificultad para conformar un grupo de expertos adecuado, tanto en cantidad como en perfil disciplinar y nivel de experiencia, especialmente en áreas técnicas o especializadas; la influencia de percepciones individuales o creencias personales que pueden comprometer la objetividad de los resultados; la falta de consenso entre expertos que conduce a resultados inconsistentes; la limitada reproducibilidad de los juicios emitidos; y la posibilidad de que los expertos se vean influidos por intereses personales o institucionales. Además, el proceso mismo implica tiempos prolongados, altos costos y condiciones logísticas complejas. Estas dificultades restringen su aplicabilidad, particularmente en contextos donde los recursos son limitados o donde se busca asegurar la replicabilidad del modelo (Odu, 2019), (Keshavarz-Ghorabae et al., 2021), (Şahin, 2021).

Como respuesta, han emergido enfoques que proponen la asignación objetiva de pesos a partir del análisis de datos, mediante algún método de ponderación de variables, con el propósito de minimizar la influencia del juicio humano. Estos métodos buscan reducir la subjetividad, así como mejorar la transparencia, consistencia y reproducibilidad del proceso de toma de decisiones. Sin embargo, también enfrentan desafíos importantes, como el tratamiento de variables heterogéneas mediante procesos de normalización y el riesgo de pérdida de información relevante durante la transformación de los datos.

Ante este panorama, se requiere el desarrollo de un modelo matemático multicriterio que integre técnicas objetivas de asignación de pesos, representando de forma estructurada y reproducible la complejidad del problema analizado. Este modelo busca reducir la subjetividad y facilitar la evaluación y selección de alternativas en escenarios reales, especialmente donde el acceso a expertos es limitado o los recursos son restringidos.

Esta investigación se aplicará en un caso de estudio real sobre accesibilidad a servicios, permitiendo validar el modelo propuesto y evaluar su utilidad práctica.

1.4. Objetivos de la investigación

En esta sección, se expone el objetivo general de la investigación, junto con los objetivos específicos que guían el desarrollo de la investigación.

1.4.1. Objetivo general

Desarrollar un modelo matemático basado en la Toma de Decisiones Multicriterio que permita asignar de manera objetiva los pesos de múltiples variables e integrarlos en una función de agregación para evaluar y seleccionar alternativas, reduciendo la influencia de juicios subjetivos.

1.4.2. Objetivos específicos

- Construir un conjunto de datos provenientes de un caso de estudio real, integrando variables heterogéneas que sustenten el análisis multivariado necesario.
- Identificar las variables significativas mediante técnicas de análisis multivariado, con el fin de estructurar un modelo matemático multicriterio que represente adecuadamente la complejidad del problema bajo estudio
- Diseñar un método de ponderación objetiva para la asignación de pesos a las variables, basado en el análisis de datos, que minimice la influencia de juicios subjetivos.
- Formular un modelo matemático multicriterio que integre los pesos asignados y una función de agregación matemática para evaluar y seleccionar las alternativas.
- Interpretar los resultados del modelo multicriterio utilizando herramientas analíticas que permitan comprender el comportamiento del problema evaluado, con el fin de fundamentar la selección de alternativas.

1.5. Hipótesis de investigación

La hipótesis de esta investigación establece:

La implementación de un modelo matemático basado en la Toma de Decisiones Multicriterio, que utilice un enfoque objetivo para la asignación de pesos a las variables, permitirá representar de manera precisa el comportamiento de un problema bajo conflicto, mejorando la evaluación estructurada y la selección fundamentada de alternativas.

1.6. Metodología de solución de la investigación

En esta investigación, se adoptó un enfoque estructurado para desarrollar un modelo matemático multicriterio que evalúe la accesibilidad a los servicios, fundamentado en la disciplina de la Toma de Decisiones Multicriterio.

La metodología propuesta consta de dos etapas. La primera se centra en la construcción de un método de ponderación de variables para la asignación de pesos; con el objetivo de determinar su importancia relativa. Estos valores se utilizarán como coeficientes en un modelo de agregación que permite formular el modelo multicriterio. Para ello, se adopta un enfoque objetivo basado en la información contenida en los datos; combinando dos técnicas complementarias para fortalecer la robustez y fiabilidad del modelo. La segunda etapa se centra en el desarrollo del modelo multicriterio, el cual permite describir, evaluar e identificar las zonas con mayor y menor accesibilidad a los servicios, considerando las distintas variables que influyen en esta problemática.

1.7. Contribuciones a las Ciencias Computacionales

- El desarrollo de un modelo matemático multicriterio que integre variables espaciales, de vivienda y educación, con el propósito de describir el comportamiento de la accesibilidad a los servicios. Este modelo proporciona una herramienta analítica que facilita la identificación y selección de las zonas de interés. A la vez que contribuye a una comprensión más profunda del problema y orienta la toma de decisiones de manera informada.
- La construcción de un método de ponderación de variables con enfoque objetivo para asignar pesos a las variables y determinar su nivel de importancia. Este método se presenta como una solución alternativa a los problemas derivados de la dependencia de expertos. La combinación de técnicas como el Análisis de Componentes Principales y el método de Entropía ayuda a mejorar la precisión en la evaluación de las alternativas. Asimismo, reduce la susceptibilidad a los efectos de la normalización de los datos. Además, al poder ser implementado en lenguajes de programación, permite realizar un análisis más eficiente y preciso en el proceso de toma de decisiones.
- La implementación de diversas técnicas matemáticas para la identificación de variables relevantes representa una contribución significativa al proceso de selección de variables. Estas técnicas, al ofrecer distintas perspectivas, permiten eliminar variables redundantes, analizar la interrelación entre las variables agrupar variables relacionadas y abordar la variabilidad de los datos; lo que mejora la precisión y la robustez de los resultados obtenidos.
- Desarrollo de un conjunto de datos integrado que incluye los valores de las variables correspondientes a las distintas zonas (localidades), su información geoespacial, los resultados de la evaluación del modelo multicriterio y el nivel de accesibilidad a las escuelas; permitiendo identificar áreas con diferentes niveles de accesibilidad y priorizar zonas de interés.

-
- Desarrollo de un conjunto de datos integral que incluye los valores de las variables relacionadas con la accesibilidad en diversas zonas. Este conjunto de datos no solo facilita la identificación de áreas con distintos niveles de accesibilidad; sino que proporcionando una base sólida para ser utilizado en futuros estudios.
 - Desarrollo de una metodología de solución aplicada al análisis de accesibilidad, que puede ser adaptada y empleada en otros contextos relacionados con la evaluación de accesibilidad a diferentes servicios. Esta metodología, basada en un enfoque multicriterio y en el análisis de datos objetivo, ofrece un marco estructurado que facilita la evaluación de la accesibilidad en diversas situaciones y permite la personalización de las variables según las necesidades de cada estudio.

1.8. Organización de la tesis

El documento de tesis está organizado de la siguiente manera:

El capítulo 2, Estado del arte, presenta una revisión de la literatura relevante en torno al problema de estudio; considerando antecedentes relacionados con la accesibilidad a servicios, los modelos matemáticos aplicados a este ámbito, los métodos y modelos de Toma de Decisiones Multicriterio, y finalmente un análisis de la bibliografía revisada.

El capítulo 3, Marco teórico, expone los conceptos fundamentales de las herramientas matemáticas y computacionales utilizadas en la investigación; incluyendo la Toma de Decisiones Multicriterio, técnicas para la asignación de pesos en modelos multicriterio, las técnicas para la identificación de variables relevantes, el análisis de accesibilidad a servicios, y en los sistemas de información geográfica.

El capítulo 4 se centra en dos puntos. Uno de ellos es la construcción de un método de ponderación de variables para la asignación de pesos, detallando el proceso matemático y la justificación del método propuesto, a fin de ser implementado dentro del marco metodológico de la Toma de Decisiones Multicriterio para obtener un modelo matemático. El siguiente punto está enfocado en la metodología implementada para el desarrollo de un modelo matemático multicriterio enfocado en la accesibilidad a servicios; junto con los hallazgos obtenidos a partir de su aplicación.

El capítulo 5 describe el proceso de validación del modelo de accesibilidad obtenido, abarcando diversos enfoques, como la comparación con modelos existentes y el análisis de sensibilidad. Además, se proporciona una interpretación del modelo en el contexto de la accesibilidad, y se incluye su representación gráfica, lo que facilita la comprensión y visualización de los resultados.

Finalmente, el capítulo 6 presenta las conclusiones de la investigación, resumiendo los principales aportes realizados, las líneas de trabajos futuros y los logros alcanzados a lo largo de la investigación.

Capítulo 2. Estado del arte

2.1. Introducción

Este capítulo ofrece un análisis exhaustivo del estado del arte, estructurado en secciones que abordan una introducción, los antecedentes históricos enfocados a la modelación y análisis de la accesibilidad, estudios recientes de modelos de accesibilidad, modelos y métodos de accesibilidad mediante la Toma de Decisiones Multicriterio. Asimismo, se presenta un análisis crítico del estado del arte en general.

La investigación sobre la accesibilidad a los servicios básicos, como escuelas, unidades de salud, parques, centros de entretenimiento, etc.; ha experimentado avances significativos. Inicialmente los modelos propuestos consideran factores espaciales, como la distancia. Sin embargo, los estudios han evolucionado hacia modelos más complejos que consideran la población y las condiciones socioeconómicas. Uno de los modelos que ha ganado relevancia, son los modelos de Toma de Decisiones Multicriterio; permitiendo integrar múltiples variables en su análisis, combinando con la tecnología de los Sistemas de Información Geográfica. Este capítulo revisa los avances históricos y recientes en el análisis, modelación y evaluación de la accesibilidad a servicios;}, destacando los modelos y métodos aplicados. Finalmente se presenta un análisis crítico sobre las fortalezas y limitaciones del estado del arte.

2.2. Antecedentes de los modelos de accesibilidad a servicios

- *Cómo la accesibilidad influye en el uso del suelo* (Hansen, 1959).

Esta investigación es uno de los primeros estudios que abordan la accesibilidad como un factor clave en el análisis de uso del suelo. Se centra en el desarrollo de un modelo que examina el crecimiento poblacional en pequeñas áreas de una región metropolitana. En particular, explora cómo la accesibilidad influye en el desarrollo residencial; considerando la interacción entre zonas comerciales y residenciales, así como la disponibilidad de espacio. El modelo se basa en la proximidad a lugares de trabajo, las oportunidades de compra y las actividades sociales, factores que determinan las decisiones sobre el uso del suelo y el desarrollo urbano.

En este documento, la accesibilidad se define como el potencial de acceder a oportunidades de interacción, ampliando el concepto tradicional que solo considera la facilidad para desplazarse de un lugar a otro. Esta definición incorpora tanto la facilidad de interacción como la cantidad de oportunidades disponibles; lo que permite una comprensión más integral. Así, la accesibilidad se interpreta como una medida de la distribución espacial de actividades en torno a un punto, ajustada a la capacidad y disposición de las personas o empresas para superar la separación geográfica.

Para evaluar el modelo propuesto, se realizó un análisis empírico utilizando datos del área metropolitana de Washington, D.C. Este análisis se basó en estudios de origen y destino realizados en 1948 y repetidos en 1955. Las variables utilizadas fueron el número de trabajos, población, tiempo de viaje, distancia, el valor del exponente y tamaño de la actividad. Se aplicó una medida de accesibilidad para calcular la variación en la accesibilidad entre áreas; es decir, una medida relativa de la

accesibilidad en la Zona 1 a una actividad ubicada dentro de la Zona 2. En los experimentos, se calcularon las mediciones de accesibilidad y se determinó el patrón de desarrollo residencial, comercial e industrial durante un período de siete años.

- *El concepto de accesibilidad: una búsqueda de una forma operativa* (Ingram, 1971).

La contribución de este trabajo se centra desarrollo de métodos cuantitativos para analizar cómo la accesibilidad influye en patrones de uso del suelo, movilidad y planificación urbana. Esta investigación propone enfoques para vincular el análisis de la accesibilidad con variables espaciales y de interacción humana; como es la accesibilidad relativa y la accesibilidad integral. En este trabajo, la accesibilidad relativa se define como el grado en que dos lugares dentro de una misma superficie están conectados. Por otro lado, la accesibilidad integral se refiere al grado de interconexión de un punto con todos los demás puntos dentro de una superficie. Esta última puede considerarse una derivación del conjunto de accesibilidades relativas desde un punto específico.

Las medidas de accesibilidad utilizadas se aplicaron en el área urbana de Hamilton, Ontario, utilizando datos de 466 áreas de la ciudad. Las medidas implementadas tomaron en cuenta el punto central de cada área, aunque en algunos casos puede haber un sesgo hacia las zonas más pobladas. Como resultado se recomienda utilizar una medida basada en la función gaussiana; ya que resulta ser más adecuada para determinar la accesibilidad integral en un punto dado.

- *Accesibilidad física como indicador social* (Wachs & Kumagai, 1973).

En esta investigación se presenta un marco conceptual para medir la accesibilidad en términos de la facilidad con la que los ciudadanos pueden acceder a diversas oportunidades de empleo y servicios. La medida propuesta se enfocó en el desarrollo de un indicador social que evalúa dicha accesibilidad. Este trabajo también se enfocó en refinar la definición de accesibilidad que relaciona la facilidad y el costo del desplazamiento de un punto a otro; proporcionando una comprensión más concreta y detallada de cómo la accesibilidad afecta la interacción con las oportunidades disponibles.

Como caso de estudio, se evaluó la accesibilidad al empleo en un condado de Los Ángeles. La zona fue subdividida en diferentes áreas y se calculó el índice de accesibilidad para cada una de ellas. Este cálculo tomó en cuenta la cantidad de sitios de empleo accesibles, el tipo de ocupación disponible y el tiempo de viaje en minutos desde cualquier ubicación. El indicador desarrollado es una suma ponderada del número de oportunidades de empleo y el tipo de ocupación que se encuentran dentro de un determinado tiempo de viaje desde alguna zona. Los pesos y los tipos de empleo considerados se determinaron en función de su relevancia e interés en el contexto del estudio.

-
- *Medición de la accesibilidad: algunos resultados preliminares* (Scientific & Company, 1976).

Este trabajo presenta un estudio centrado en medir la accesibilidad a las zonas de empleo en una ciudad, utilizando como base la medida propuesta por (Hansen, 1959). Los datos utilizados provienen de la Encuesta de Viajes. El estudio para el desarrollo del índice de accesibilidad incluye variables como el empleo, los hogares, la población y la modalidad de transporte en automóvil.

Un aspecto clave de este estudio fue el análisis de sensibilidad de la accesibilidad. Este análisis examinó cómo los cambios en los valores de los parámetros afectan la estabilidad del índice ante variaciones en el contexto urbano. Las variaciones en los parámetros se reflejaron en los efectos de la configuración zonal y la agregación zonal. Como resultado, se observó que el índice de accesibilidad mostró sensibilidad a los ajustes realizados, lo que destaca la importancia de considerar estos factores al medir la accesibilidad en entornos urbanos.

- *La ecología intraurbana de la atención médica primaria: patrones de accesibilidad y sus implicaciones políticas* (Knox, 1978).

Esta investigación examina los patrones de accesibilidad intraurbana para la atención médica primaria en cuatro ciudades de Escocia, adoptando un enfoque centrado en las políticas públicas y el bienestar comunitario. Se emplea un modelo de interacción modificado para analizar la accesibilidad local a los centros de atención primaria. La información sobre la ubicación de los centros de atención médica fue obtenida de listas médicas proporcionadas por las juntas de salud. Además, las subdivisiones ecológicas de cada ciudad se establecieron a partir de un análisis de conglomerados.

El análisis de la accesibilidad intraurbana a la atención médica primaria se basó en un modelo de gravedad (interacción). Este modelo considera factores como el tamaño de las instalaciones quirúrgicas en un barrio específico, la distancia lineal entre dos barrios, y una función de disminución de la distancia para el registro de pacientes con médicos de familia. Un aspecto distintivo de este modelo es su consideración de la modalidad de transporte, ya sea automóvil privado o transporte público. Los resultados muestran las zonas más desfavorecidas de cada ciudad con respecto a la escasez de médicos. También se han detectado patrones espaciales que revelan la existencia de desventajas socioeconómicas en las zonas analizadas.

Finalmente, Tabla 2.1 presenta un resumen de los trabajos relacionados con los antecedentes de los modelos de accesibilidad, abarcando los modelos y variables utilizadas en el contexto de análisis.

- *Un indicador de accesibilidad a unidades de servicios clave para ciudades mexicanas: fundamentos, diseño y aplicación* (Garrocho, C. F. & Campos, 2006).

En este trabajo se diseñó un índice de accesibilidad a los servicios con el objeto de contar con un indicador que apoye la toma de decisiones de los planificadores y administradores urbanos en el contexto de México. El índice de tipo interacción espacial combina la disponibilidad con los costos de transporte, su relación con la oferta del servicio; además considera un parámetro que se ajusta ante los cambios de

los costos de transporte y representa qué tan sensible es la demanda. Para probar la capacidad del índice de accesibilidad a unidades de servicios, como caso de estudio se consideraron los servicios de salud pública para la población mayor de 65 años sin seguridad social del área metropolitana de Toluca, Estado de México.

Los datos recolectados se obtuvieron a partir de las estimaciones del Instituto de Salud del Estado de México, como la capacidad de atención de cada unidad de salud y la población que ganan menos de dos salarios mayor a 65 años sin seguridad social. Las distancias se calcularon mediante el Sistema de Información Geográfica ArcView, considerando desde donde radica la población hacia la ubicación de la unidad de salud. Los resultados del índice de accesibilidad fueron mostrados a través de un mapa, el cual permite observar las desigualdades en el acceso. Posteriormente se realizó la simulación de escenarios para evaluar los posibles efectos que tienen algunas acciones de política pública y analizar el impacto en términos de la accesibilidad a los servicios de salud. Por tal motivo al índice de accesibilidad considera elementos importantes que afectan la accesibilidad que ayuda a la toma de decisiones.

La Tabla 2.1 muestra un resumen de la revisión bibliográfica realizada, correspondiente a los antecedentes del estudio.

Tabla 2.1. Resumen de los antecedentes de la accesibilidad

Trabajo	Modelos, métodos y medidas de accesibilidad	Variables	Contexto de la accesibilidad
(Hansen, 1959)	Medida de accesibilidad relativa	población, tiempo de viaje, distancia	Proximidad a las actividades sociales
(Ingram, 1971)	Medida de accesibilidad integral y accesibilidad relativa	Área, distancia, función de atracción: normal, recíproca, exponencial negativa y rectangular	Acceso entre áreas de una ciudad.
(Wachs & Kumagai, 1973)	Indicador de accesibilidad	Número de sitios de empleo, radio del tiempo de viaje, tipo de empleo, tiempo de viaje	Empleos
(Scientific & Company, 1976)	Medida de accesibilidad relativa	Zonas de empleo, hogares, modalidad transporte público	Zonas de empleo
(Knox, 1978)	Modelo de interacción	Ubicación de los centros médicos, distancia, tamaño de la instalación, modalidad de transporte automóvil y transporte público, función de distribución	Atención médica
(Garrocho, C. F.; Campos, 2006)	Índice de accesibilidad a unidades de servicio (índice de interacción espacial)	Capacidad de atención, población potencial, distancias,	Unidades de salud

2.3. Modelos y métodos de accesibilidad a servicios

- *Accesibilidad basada en un modelo de gravedad-radiación y una API Google Maps: un caso de estudio en Australia* (Xia et al., 2018).

En este estudio se construyó un modelo integrado de Gravedad-Radiación para medir la accesibilidad a las principales ciudades. También se propuso un marco para adquirir datos a partir de una API con mapeo web. Posteriormente se midió la accesibilidad general a una escala de ciudad, estado y nacional. El propósito del modelo integrado de Gravedad-Radiación es estimar la probabilidad de viaje y la población potencial que viaja (EPTP). El modelo de gravedad se utilizó para determinar la capacidad de atracción de diferentes áreas urbanas importantes (SUA) en la red. Los datos obtenidos están basados en los resultados del Censo de Población y Alojamiento, los cuales muestran celdas de cuadrícula de 1 km × 1 km en toda Australia. Estos datos fueron introducidos en la cuadrícula del nuevo modelo de Gravedad-Radiación para estimar la probabilidad de viaje y la población potencial que viaja (EPTP). Asimismo, se utilizó la API de Google Maps para adquirir costos de viaje a nivel de cuadrícula. Posteriormente se implementaron indicadores ponderados para evaluar la accesibilidad. Para la validación de los datos se analizaron a diferente escala, cuyo propósito fue medir la confiabilidad de la información. Los resultados muestran que la EPTP está fuertemente correlacionada con la población residente tanto a nivel SUA como a nivel estatal en Australia.

- *Diseño metodológico para estimar indicadores de accesibilidad en entornos periféricos de una zona metropolitana* (Biosca & Espinosa, 2018).

La investigación se centra en determinar la accesibilidad a las periferias en ámbitos urbanos. Este estudio propone un diseño metodológico que incorpora el indicador de separación y el indicador de interacción espacial para analizar la accesibilidad mediante el sistema de transporte, con un enfoque en la Zona Metropolitana de Querétaro como caso de estudio.

Los datos del tiempo de viaje se obtuvieron mediante un sistema de posicionamiento global en vehículos. A partir de la información de distancia y tiempos de viaje, se aplicó el algoritmo de Dijkstra para determinar la ruta más corta en términos de tiempo y distancia. Además, se estimó el valor del tiempo relacionado con motivos laborales utilizando fuentes oficiales. Se calcularon también los costos de transporte considerando los precios de la gasolina y el rendimiento promedio. Para el parámetro de fricción de la distancia, se empleó un modelo gravitacional. La modalidad de transporte considerada fue autobús y automóvil. Todos estos datos se utilizaron en los indicadores de separación e interacción espacial.

Los resultados muestran que existe una relación directa entre la accesibilidad a las áreas periferias con la tasa media de viajes en autobús. Por lo que dichos resultados, muestran que tanto el indicador de separación como el de interacción espacial corroboran que la variable de accesibilidad está ligada a la elección del modo de transporte público, específicamente el autobús.

-
- *Evaluación de la accesibilidad urbana: aprovechamiento datos de código abierto y analítica para superar limitaciones existentes* (Logan et al., 2019)

El objetivo de esta investigación es proponer una metodología para medir la proximidad a servicios, en términos de la accesibilidad, tomando en cuenta los avances tecnológicos enfocado a los datos abiertos y el análisis de código abierto.

En el caso de estudio se consideraron tres ciudades de EE. UU.: Chicago, Baltimore y Detroit. La modalidad de transporte se consideró caminando, bicicleta para ir a las escuelas secundarias públicas, atención médica y parques. Los datos se obtuvieron principalmente de OpenStreetMap (OSM). Posteriormente se empleó un algoritmo de enrutamiento de código abierto para determinar las distancias y tiempos a los destinos más cercanos. Asimismo, se realizó la asignación de información demográfica a unidades de resolución más finas, donde la población del bloque del censo se dividió por igual entre los edificios restantes y se excluyeron los edificios que no están zonificados. A este proceso se le llama mapeo dasimétrico.

En este enfoque la accesibilidad mide la distancia de la red y el tiempo desde cada edificio hasta el punto de destino más cercano, a fin de poder realizarlo a gran escala y promover el desarrollo de código abierto para que otros puedan llevar a cabo el mismo análisis. También se consideraron umbrales de distancia. Se puede observar que este análisis de resolución fina a gran escala puede ser estudiado para grandes ciudades enfocado en la proximidad

- *Análisis de accesibilidad espacial de los centros de servicios sanitarios de Gamo Zona Gofa, Etiopía a través de la técnica geoespacial* (Dejen et al., 2019).

El propósito de esta investigación es analizar la distribución y accesibilidad espacial de los centros de servicios de salud mediante Sistemas de Información Geográfica de la Zona de Gamo Gofa de la región Pueblos de Naciones y Nacionalidades del Sur (SNNP), Etiopía.

Los datos extraídos para el estudio consideraron la cobertura del suelo, el modelo de elevación digital, la pendiente de la calle, la distribución de la población, el acceso por la red vial, la ubicación geográfica de los servicios de salud y el escenario de viaje. Los datos de cobertura del suelo se produjeron a partir de una imagen mediante Landsat TM. Los datos fueron extraídos de Google Earth. La ubicación de los centros de servicios se obtuvo mediante GPS. Asimismo, se tomó una muestra de los conductores de automóviles y se realizó una entrevista para obtener los datos sobre el tiempo de viaje, velocidad y la modalidad de transporte. Para la accesibilidad de los servicios de salud, se utilizó ArcGis 9.3.1 y ERDAS IMAGINE 9.1. para crear mosaicos de imágenes satelitales y terrestres, específicamente se ocupó AccessMod 4.0

En este estudio se tomaron en cuenta cuatro escenarios para estudiar la accesibilidad física, considerando la modalidad de transporte, tipo de centros de salud de acuerdo a las normas de la Organización Mundial de la Salud.

- o Tiempo de viaje de 30 y 60 minutos a los centros de salud en la modalidad a pie y transporte motorizado

-
- o El tiempo de viaje considerado fue de 30 y 60 minutos a hospitales en la modalidad a pie y transporte motorizado.

En el resultado se pudo observar que tomando en cuenta la modalidad de transporte motorizado, aumentó la accesibilidad a los centros de servicios de salud; a diferencia de la modalidad a pie. El uso de la tecnología geoespacial mostró ser una herramienta eficaz y robusta en la investigación relacionado con aspectos de la accesibilidad a la atención médica.

- *Accesibilidad a equipamientos según movilidad y modos de transporte en una ciudad media, Los Ángeles, Chile* (Quezada et al., 2019).

Este trabajo tiene como objetivo analizar accesibilidad a diversos equipamientos en la Ciudad de los Ángeles, Chile. Los equipamientos que se consideraron fueron educación, salud, servicios y comerciales en sus diversos niveles; por ejemplo, las clínicas de salud, hospitales, centro de salud, establecimientos públicos y privados de educación, mercados, farmacias, perfumerías, etc. Los modos de transporte que se tomaron en cuenta fueron el transporte privado, público y caminata.

En el análisis de accesibilidad se contemplaron diversas medidas: Indicador de Tiempos de viaje a Educación y Salud. Este indicador se utilizó para estimar el tiempo de viaje desde un hogar hacia el equipamiento, empleando el método de interpolación Inverse Distance Weighting (IDW) con el uso de la herramienta Spatial Analyst de ArcGis 10.1. Otro indicador que se utilizó fue el de espacio de actividad o áreas de servicios del comercio. Este indicador se define por un área buffer de cobertura, estimada por la configuración espacial de los puntos de oferta y la velocidad de desplazamiento por la red en la modalidad caminando. Por último, se aplicó el indicador de oportunidades acumuladas en la modalidad de caminata. Este indicador mide los niveles de accesibilidad, calculando el número de equipamientos comerciales y de servicios, tomando en cuenta el tiempo de viaje.

El estudio realizado ayudó a identificar aquellas zonas más y menos favorecidas con respecto a la accesibilidad. Lo anterior fue debido a los métodos empleados, ya que este enfoque combina el análisis de redes y la movilidad de los modos de transporte desde una visión espacial y social.

- *Accesibilidad a oportunidades a partir de datos de transporte público monitoreados por GPS: El caso de Santiago de Chile* (Basso et al., 2020).

El estudio analiza la accesibilidad del transporte público en Santiago de Chile a la educación, la salud y las oportunidades laborales. Para la obtención de los datos se dividió la ciudad en 1390 zonas transitables utilizando un sistema geodésico de cuadrícula global discreta basado en hexágonos. Luego se calcularon los tiempos de viaje entre estas zonas durante una semana utilizando a partir de GPS y tomando en cuenta 6681 autobuses del sistema de transporte público. A partir de estos datos, se determinó el número de oportunidades alcanzables desde cada zona al resto de zonas, agregándolas a nivel municipal. El indicador de accesibilidad propuesto, basado en trabajos previos, cuenta el número de oportunidades que se pueden alcanzar en un tiempo de viaje determinado, lo que se conoce como una medida de contorno general.

Los resultados muestran que el transporte público no atiende por completo la desigualdad en toda la ciudad, dada la distribución geográfica de las oportunidades; ya que existen algunos municipios de menores ingresos en el sur quedan excluidos de empleos y servicios de salud. Esto resalta la desigualdad social que existen en las personas de bajos ingresos, quienes suelen tener más limitaciones en su ubicación y elección de modo de transporte.

- *La organización espacial de la accesibilidad y la jerarquía funcional: el caso de Israel* (Kaplan et al., 2020).

El objetivo de esta investigación es analizar la accesibilidad espacial de la vía nacional multiescalar; con el fin de captar distintos sistemas de accesibilidad espacial y la relación con la formación de sistemas funcionales, desde un nivel local hasta nacional. La metodología propuesta consistió de cuatro aspectos: la caracterización de la estructura jerárquica de los sistemas de accesibilidad espacial, la identificación de sus estructuras funcionales; examinación de las asociaciones entre sistemas espaciales y funcionales en las diferentes escalas; finalmente, la detección de las transiciones entre los sistemas funciones en términos de las escalas espaciales.

El caso de estudio se enfocó en Israel. En los modos de viaje de la red vial se consideró auto privado, autobús público, a pie, tren y otros. Los datos de la red de carreteras se extrajeron de la red nacional de carreteras. A los datos recolectados, se le simplificaron sus geometrías para reducir complejidad, se tomaron en cuenta los caminos pavimentados, entre otros análisis. El procesamiento de los datos se realizó con ArcGis. Para evaluar la accesibilidad se tomó en cuenta una medida de centralidad de integración que corresponde a la medida de centralidad de cercanía basada en gráficos. Esta medida define que tan cerca está un nodo, definido como segmento, de todos los demás nodos; el cual es representado como el grado de accesibilidad para cada segmento de la red (carretera). Las medidas se calcularon para un gran conjunto de radios que representan diferentes escalas espaciales. Asimismo, se utilizó el Análisis de Componentes Principales (PCA) y la Factorización del eje principal (PAF) para identificar variaciones significativas en las tendencias de flexibilidad en la distribución espacial y para extraer la estructura de la centralidad latente. Los resultados mostraron tres sistemas de accesibilidad espacial: local, regional y nacional; para este proceso se utilizó SPSS.

- *Eficiencia y equidad espacial como principios para la planificación territorial de los establecimientos educativos: análisis de la educación primaria al sureste del municipio Maracaibo, Venezuela* (Montes et al., 2020).

Esta investigación se centró en estudiar la accesibilidad a través de la eficiencia y equidad espacial en la distribución de establecimientos educativos de educación primaria en el sureste del municipio Maracaibo, Venezuela. Su objetivo principal fue evaluar la situación proponer mejoras mediante un modelo de localización-asignación (MEDIRES). Este modelo optimiza la ubicación de nuevos establecimientos educativos para reducir las desigualdades en el acceso a los servicios educativos, mejorando tanto la eficiencia como la equidad espacial.

La metodología se enfocó en la construcción de una base de datos geográfica a partir de encuestas para comprender las preferencias de traslado de la población y la

digitalización de la infraestructura vial y la ubicación de la demanda población. A partir de lo anterior, se integraron técnicas de análisis espacial, sistemas de información geográfica y la construcción del modelo p-mediana con restricciones de distancia para minimizar los recorridos totales y equilibrar la distribución espacial (MEDIRES). Las variables consideradas incluyeron distancia y variabilidad en las distancias, oferta educativa, demanda poblacional, red vial, costos de transporte y distancias óptimas.

Como resultado, la propuesta redujo significativamente la distancia total recorrida y aumentó la proporción de población servida dentro del rango ideal de distancia a los establecimientos, con mejoras del 47,07% en eficiencia y del 65,52% en equidad. Este trabajo ofrece una solución práctica para optimizar la planificación territorial en servicios educativos, destacando su potencial para promover un desarrollo más inclusivo y sostenible.

- *Método jerárquico de área de captación flotante de dos pasos (2SFCA): medición de la accesibilidad espacial a instalaciones de atención médica jerárquicas en Shenzhen, China* (Tao et al., 2020).

Este trabajo se enfocó en analizar la accesibilidad a las instalaciones de atención médica en Shenzhen, China con el propósito de identificar las disparidades espaciales en la accesibilidad y proponer un marco metodológico para optimizar la distribución de los recursos de atención médica. Este estudio se centró en la introducción de un método jerárquico de dos pasos (H2SFCA por sus siglas en inglés) que mejora las metodologías tradicionales al considerar diferentes niveles de servicios médicos y jerarquías de atención.

La metodología se enfocó en la recolección de los datos enfocados en la capacidad de las instalaciones, la demanda de la población, la distancia y los niveles de atención médica. A partir de lo anterior se realizó una jerarquización de las instalaciones médicas con base al nivel de atención. Posteriormente, se empleó método de área de captación flotante de dos pasos (2SFCA por sus siglas en inglés) y modelo jerárquico de accesibilidad (H2SFCA) para evaluar la accesibilidad a los diferentes niveles de atención médica. Asimismo, se utilizaron sistemas de información geográfica para la representación gráfica y la evaluación espacial.

Los resultados mostraron disparidades significativas en la accesibilidad, especialmente en las áreas periféricas, que tienen menor acceso en comparación con las zonas centrales. El modelo H2SFCA resultó ser una herramienta efectiva para identificar desigualdades espaciales y proponer estrategias de redistribución de recursos que mejoran la equidad y accesibilidad en áreas vulnerables.

- *La última milla en el acceso a la escuela: mapeo de los desiertos educativos en los países en desarrollo* (Rodríguez-Segura & Kim, 2021).

El propósito de esta investigación es identificar sistemáticamente zonas, denominadas “desiertos educativos”; para conocer el alcance físico hacia las escuelas en las poblaciones de un país. La metodología principalmente consiste en un algoritmo de código abierto para identificar con precisión las áreas pobladas que no tienen acceso físico a las escuelas. Como caso de estudio se consideró el país de Guatemala.

Los datos extraídos están relacionados con la ubicación geográfica de las escuelas y la distribución geográfica de la población. Posteriormente se calculó la distancia en línea recta entre el centro de cada unidad de población y su escuela pública más cercana. En este estudio se estimó la distancia que debe recorrer una persona para llegar a una escuela. También se estableció un umbral de referencia con base a las normas de accesibilidad y eficiencia determinadas. Los datos de la población se trabajaron con valores de grano fino y geolocalizados del WorldPop, donde se visualizaron las capas que proporcionan estimaciones de la distribución de la población a una resolución de aproximadamente mil metros cuadrados para identificar más a fondo las áreas que no tienen acceso hacia alguna escuela. Como resultado se obtuvieron conjuntos de terrenos geolocalizados con información de las variables del valor estimado población que habita en cada parcela, y la distancia mínima de dicha parcela a una pública escuela primaria.

- *Evaluación de la accesibilidad a los ASF desde las paradas de autobús utilizando medidas de distancia: caso de dos ciudades indias* (Yenisetty & Bahadure, 2021).

El objetivo de esta investigación es evaluar la accesibilidad a amenidades, servicios e instalaciones (amenities, services, and facilities, ASF por sus siglas en inglés) basado en medidas de distancia. Y mediante la conexión del sistema de transporte público. Como caso de estudio se tomaron dos ciudades de la India, Nagpur y Bhopal. Los datos extraídos fueron recolectados de diversas fuentes; por ejemplo, la plataforma de Datos abiertos del gobierno de la India, la Oficina del ReSIGtrador General y Comisionado del Censo, Servicio Geológico, Mapas de uso del suelo, entre otros. Los variables considerados fueron la población, longitud media del viaje, pendiente de área, ubicación de los servicios (hospitales, escuelas, parques, estación de policía, oficina de correos, salones comunitarios, la infraestructura vial, entre otros.

Para el análisis se realizó un método basado en cuadrículas hexagonales sobre el mapa estimar la población en la ciudad. La distancia obtenida consideró desde la parada de autobús más cercana de las cuadrículas hexagonales hacia los sectores de ASF. Para la obtención del camino más corto se utilizó algoritmo de análisis de redes Dijkstra. Para el cálculo de la distancia, se consideraron valores hasta los 2400 m; para valores mayores a este rango de distancia, los servicios de transporte público se toman como no atendidos. Asimismo, el análisis fue mediante la distancia euclidiana. El procesamiento de los datos se realizó mediante el Sistema de Información Geográfica ArcGis 10.5.

A partir de lo anterior, se aplicó un análisis estadístico para analizar la correlación entre las distancias de las ciudades consideradas en el caso de estudio; también se tomaron valores de R^2 . Para ello se utilizó el software SPSS 16.0. Posteriormente se tomaron en cuenta diversos rangos de distancia para clasificarlos en colores y poder medir la accesibilidad. Los resultados muestran el nivel de accesibilidad a través de porcentaje de distancia euclidiana por sectores de ASF para las ciudades de Nagpur y Bhopal.

- *Métodos de medición de la accesibilidad espacial para el sistema de salud in Uganda* (Ouma et al., 2021).

Este trabajo tiene como propósito estimar el acceso espacial a la atención médica para el caso de estudio de Uganda. Para modelar la accesibilidad espacial se obtuvieron

datos de portal del RCMRD. Las carreteras se obtuvieron usando *OpenStreetMaps* y *Google Map Maker*. Se recopilaron datos de los cuerpos de agua y ríos de Global Lakes y de base de datos sobre humedales. También se descargó un modelo de superficie ráster de elevación digital (DEM) de la misión de topografía de radar *Shuttle* de la NASA en el sitio web del Centro de Archivo Activo Distribuido de Procesos Terrestres (LP DAAC).

Para la distribución de la población se tomó un conjunto de datos en cuadrícula con una resolución de 0.1×0.1 km. La modalidad de transporte considerada fue caminar, andar en bicicleta y el tiempo de viaje motorizado. En la investigación se utilizó un análisis de costo-distancia para modelar el tiempo de viaje para estimar el tiempo de viaje al centro de atención primaria de salud más cercano. El modelo se ajustó a las diferencias en el uso de la tierra, los patrones climáticos y la elevación del suelo; al mismo tiempo el modelo excluye barreras como cuerpos de agua otras áreas. En esta investigación fue necesario considerar los efectos que provocan el uso de los caminos cuando llega la época de lluvias, ya que las carreteras pueden ser intransitables. Los resultados permitieron proporcionar información útil sobre dónde se encuentran las brechas en la accesibilidad.

- *Acceso a parques urbanos: Comparación de medidas de accesibilidad espacial utilizando tres enfoques basados en SIG* (S. Wang et al., 2021).

Esta investigación toma en cuenta diferentes métodos para medir la accesibilidad a parques urbanos mediante Sistemas de Información Geográfica. El estudio se centra en la accesibilidad a parques urbanos en dos ciudades con características contrastantes: Ipswich, Australia, una ciudad extensa con más de 500 parques distribuidos en 1,090 km²; y Enschede, Países Bajos, una ciudad compacta con alta accesibilidad peatonal y ciclista. Se comparan métodos en estos diferentes contextos urbanos para evaluar su aplicabilidad en distintos entornos geográficos.

Los datos utilizados provienen del Portal de Datos Abiertos del Gobierno de Queensland, OpenStreetMap y censos nacionales. Se incorporaron variables como la ubicación de los parques, su tamaño, densidad poblacional, distancias de viaje a pie y en vehículo, patrones de urbanización, la red vial y los umbrales de distancia.

El propósito de esta investigación consiste en determinar qué enfoques proporcionan medidas más precisas y útiles para la planificación urbana. La principal contribución del artículo es la evaluación de tres métodos de medición de accesibilidad a parques urbanos: un índice estadístico que mide la disponibilidad de parques por unidad geográfica; un modelo de proximidad espacial, que evalúa la accesibilidad en función de la distancia a los parques más cercanos, considerando la red vial y diferentes umbrales de distancia; y finalmente un modelo de interacción espacial, que pondera la accesibilidad en función de la atracción de los parques según su tamaño y la densidad de población.

El análisis reveló que el umbral de distancia y el modo de transporte influyen más en la accesibilidad que la elección del destino, destacando la importancia de considerar la conectividad de la red vial y los accesos reales a los parques. El estudio concluye que no existe un único método óptimo para medir la accesibilidad, sino que la elección del modelo depende del contexto urbano.

-
- *Evaluación de la accesibilidad urbana en Monterrey, México: un enfoque transferible para evaluar el acceso a los principales destinos a nivel metropolitano y local* (Gaxiola et al., 2021).

Este trabajo tiene el propósito de determinar la accesibilidad de la zona metropolitana de la Ciudad de Monterrey (MMZ) y la zona local denominada Distrito Tec (zonas aledañas al campus universitario del Tecnológico de Monterrey). Para evaluar la accesibilidad se utilizó la herramienta de código abierto Urban Mobility Accessibility Computer (UrMoAC), la cual fue desarrollada por el Centro Aeroespacial Alemán (DLR). Esta herramienta permite calcular medidas de accesibilidad dentro de un área geográfica; la cual proporciona resultados precisos basados en distancias reales utilizando algún modo de transporte.

Este trabajo está basado en dos escalas. El propósito del trabajo en la zona Distrito Tec es analizar los niveles de accesibilidad, desde una perspectiva local, a los principales puestos de trabajo, escuelas, supermercados y hospitales, mediante uso de la bicicleta y a pie. El análisis se realizó a nivel de manzana, utilizando datos extraídos de las Áreas Geoestadísticas Básicas (AGEB) del Instituto Nacional de Estadística y Geografía (INEGI). Las variables utilizadas para la accesibilidad del Distrito Tec fueron el tiempo de viaje, distancia, velocidad media, modo de transporte un umbral de 15 minutos.

El estudio en la ZMM tiene el objetivo de demostrar la diversidad existente en la accesibilidad en la modalidad de automóvil y la relación con indicadores socioeconómicos. El análisis se realizó a nivel de manzana, utilizando datos extraídos de las Áreas Geoestadísticas Básicas (AGEB) del Instituto Nacional de Estadística y Geografía (INEGI). También se extrajeron datos del índice de marginación para observar cómo se encuentran las diversas zonas de ciudad. Los destinos considerados fueron las escuelas, hospitales, principales centros de empleo bicicleta y a pie. Los criterios utilizados para medir la accesibilidad fueron, distancia, tiempo y un umbral de 15 minutos.

- *Accesibilidad geográfica a la educación secundaria superior: Un estudio de caso regional en Italia* (Bruno et al., 2022).

Este trabajo analiza las desigualdades en el acceso a la educación secundaria superior en la región de Campania, Italia. El estudio se centra en la educación secundaria superior, que se divide en liceos (educación general en humanidades, ciencias y lenguas extranjeras) e institutos técnicos y vocacionales (formación profesional en tecnología, economía y servicios). La investigación abarca 541 municipios de Campania con aproximadamente 6 millones de habitantes.

Para el análisis, se emplearon datos oficiales del Ministerio de Educación de Italia, el Instituto Nacional de Estadística y la Autoridad Italiana de Comunicaciones. Se integraron variables como la ubicación de las escuelas, la distribución de la población, la infraestructura vial, los ingresos per cápita y la cobertura de internet.

El objetivo consistió evaluar cómo la ubicación geográfica influye en la equidad educativa, utilizando indicadores espaciales que permiten medir las barreras en el acceso a diferentes tipos de escuelas y programas académicos. En este estudio se aplicó utilizando un enfoque basado en Sistemas de Información Geográfica. Se

analizó tanto la proximidad a las escuelas como la conectividad a la red educativa, en función del tiempo de viaje y la variedad de programas disponibles.

Los resultados revelan fuertes desigualdades en el acceso a la educación secundaria superior en Campania. El estudio demuestra que la ubicación geográfica es un factor determinante en la equidad educativa. Se recomienda la expansión de la red escolar en áreas con menor cobertura, la redistribución de programas académicos para garantizar mayor diversidad de opciones y la mejora en la infraestructura de transporte y conectividad digital.

- *Integración de la función de conjuntos difusos en el método de área de captación flotante para refinar la accesibilidad espacial a los puntos de servicio* (Mahuve, E. F.; Tarimo, 2022).

Este estudio se enfoca en la medición de la accesibilidad a servicios urbanos mediante la incorporación de la teoría de conjuntos difusos en el método de Área de Captación Flotante (FCA, por sus siglas en inglés). La investigación se centra en la accesibilidad a diversos puntos de servicio, incluyendo hospitales, comercios y oficinas gubernamentales, con un énfasis particular en los servicios médicos y comerciales. Para el análisis, se emplearon datos geoespaciales sobre la ubicación de los puntos de servicio, la distribución poblacional y la red vial, obtenidos de plataformas de información oficial y sistemas de información geográfica.

El objetivo es representar de manera más realista la transición gradual en el acceso a los servicios en función de la distancia. La principal contribución radica en la implementación del modelo de área de captación flotante difusa (FFCA por sus siglas en inglés); el cual incorpora funciones de pertenencia difusas para calcular la accesibilidad de manera continua. Se aplicó una función de decaimiento difusa que pondera la accesibilidad en función de la distancia real a los servicios, permitiendo una representación más detallada de la accesibilidad espacial. Este enfoque permite una evaluación más precisa de la distribución de los servicios en el espacio urbano; superando las limitaciones de los modelos tradicionales que aplican cortes abruptos en la accesibilidad.

Los resultados indican que el modelo FFCA proporciona una evaluación más precisa y equitativa de la accesibilidad, identificando con mayor claridad las áreas con acceso parcial o deficiente y evitando las distorsiones presentes en los modelos convencionales.

La integración de conjuntos difusos mejora la flexibilidad del análisis, permitiendo su adaptación a distintos tipos de servicios urbanos. El estudio demuestra que el uso de conjuntos difusos en el cálculo de accesibilidad espacial permite una representación más realista del acceso a los servicios, facilitando una mejor planificación urbana y promoviendo una distribución más equitativa de las infraestructuras esenciales.

- *Investigación sobre la accesibilidad a hospitales en regiones frías: Un estudio de caso en Harbin, China* (Wu et al., 2023).

Este estudio analiza el impacto de las condiciones climáticas extremas en la accesibilidad a los servicios de salud en Harbin, China, una ciudad caracterizada por

inviernos extremadamente fríos. Harbin tiene una población aproximada de 3.96 millones de habitantes y una extensión territorial de 591 km², distribuida en cinco distritos principales. Los datos utilizados provienen de fuentes como el Censo de 2010 y la plataforma de datos abiertos del gobierno municipal de Harbin; los cuales incluyen información sobre hospitales, distribución poblacional, red vial y patrones de movilidad urbana.

El estudio emplea Sistemas de Información Geográfica en combinación con un modelo de barrera espacial ponderado para analizar la accesibilidad hospitalaria en verano e invierno. Su principal contribución es la aplicación de este modelo, que evalúa cómo el hielo y la nieve afectan la movilidad y; en consecuencia, el acceso a los hospitales. Este enfoque permite identificar deficiencias en la distribución de hospitales y proponer estrategias para mejorar la equidad en la atención médica en ciudades con inviernos severos.

Las variables consideradas incluyen la red vial, la jerarquía de carreteras, los tiempos de viaje reales en cada estación, la ubicación de los hospitales, la densidad poblacional y la eficiencia del transporte en condiciones climáticas adversas. La accesibilidad hospitalaria se clasifica en cinco niveles, permitiendo una evaluación detallada de las áreas con mayor y menor acceso a los servicios de salud.

Los resultados muestran que la accesibilidad hospitalaria en Harbin es desigual y altamente dependiente del clima. Durante el verano, los hospitales del centro urbano presentan una accesibilidad óptima. Sin embargo, en invierno los tiempos de viaje aumentan significativamente, reduciendo la accesibilidad en las zonas periféricas.

El estudio concluye que la planificación hospitalaria en ciudades frías debe considerar el impacto del clima en la movilidad urbana, proponiendo estrategias de redistribución de hospitales, mejoras en la infraestructura vial y opciones de transporte especializado para garantizar un acceso equitativo a los servicios de salud.

- *Análisis de la desigualdad en la accesibilidad a los servicios según la ubicación del vecindario y las características socioeconómicas en Delhi* (Marwal & Silva, 2023)

El objetivo de esta investigación es evaluar las desigualdades en el acceso a servicios urbanos esenciales; como escuelas, hospitales, instalaciones de entretenimiento y centros de empleo. A través de este análisis, se busca determinar si las desigualdades en la accesibilidad están más influenciadas por la distancia geográfica o por factores como la riqueza del vecindario y la composición social. El estudio consideró Delhi, India; una de las ciudades más pobladas del mundo. Los datos obtenidos para el análisis provienen de diversas fuentes oficiales, como la dirección de educación del gobierno de Delhi, el departamento de salud y bienestar familiar, y el Plan Maestro de Delhi. También se consideraron datos demográficos y económicos de 4,145 ubicaciones residenciales, complementados con encuestas de ingresos a 1,700 hogares y mediciones de densidad poblacional.

El estudio aplicó el modelo de área de captación flotante en dos pasos mejorado (por sus siglas en inglés E-2SFCA). También, utilizó un modelo de regresión geográficamente ponderada por sus siglas en inglés GWR) para analizar cómo la accesibilidad varía en función de la ubicación y las condiciones socioeconómicas de

los vecindarios. Entre las principales variables analizadas se encuentran la distancia al centro de la ciudad, la riqueza, la proporción de población de castas registradas, la densidad de población, la ubicación espacial de los vecindarios y otras características socioeconómicas.

El estudio confirma la existencia de desigualdades espaciales en la accesibilidad a servicios urbanos en Delhi. Sin embargo, la relación entre accesibilidad y características socioeconómicas varía según el tipo de servicio. Los resultados también muestran que la ubicación geográfica es el principal factor que determina la accesibilidad a los servicios, por encima de las diferencias socioeconómicas.

- *Modelo Mejorado de área de captación virtual en dos pasos (E2SVCA) para medir la accesibilidad a telesalud* (Shao & Luo, 2023).

El artículo tiene como objetivo desarrollar un método para evaluar la accesibilidad a servicios médicos remotos mediante videoconferencia en el condado de Cook, Illinois, EE. UU. Los datos empleados para el análisis provienen de múltiples fuentes oficiales. Tales como la ubicación de médicos obtenida de Sharecare.com, la información demográfica del censo de EE. UU., los datos de la red vial son proporcionados por OpenStreetMap y procesados mediante OSRM, las velocidades de internet suministradas por la Comisión Federal de Comunicaciones y los requisitos técnicos para videoconferencias establecidos por Zoom. La zona de estudio cuenta con una población aproximada de 5.2 millones de habitantes distribuidos en 99,036 bloques censales.

La metodología empleada se basa en el modelo de área de captación virtual en dos pasos mejorado (por sus siglas en inglés E2SVCA); que adapta el método tradicional del área de captación flotante en dos pasos (2SFCA), integrando métricas específicas para evaluar la calidad del acceso a internet (velocidad mínima, máxima más frecuente y mínima más frecuente). Este modelo permite generar escenarios más precisos sobre la accesibilidad real a la telesalud. Las variables consideradas en el modelo incluyen el número de médicos disponibles, la cantidad de habitantes por bloque urbano, el tiempo estimado de viaje hacia los centros de salud, los niveles de complejidad hospitalaria (baja, media y alta) y la estratificación socioeconómica.

La contribución de este trabajo radica en el perfeccionamiento del modelo 2SVCA mediante la incorporación de una función escalonada que evalúa diferentes categorías de calidad del ancho de banda requerido para teleconsultas. Este enfoque supera la limitación previa, que únicamente consideraba la disponibilidad básica de internet. Los resultados obtenidos revelan que el modelo 2SVCA tiende a sobreestimar la accesibilidad al considerar solo la existencia de internet, sin evaluar adecuadamente su calidad. Por el contrario, el modelo E2SVCA identifica claramente áreas con accesibilidad limitada debido a velocidades insuficientes, demostrando la importancia de evaluar la calidad efectiva del internet para realizar una medición precisa de la accesibilidad a la telesalud.

-
- *Análisis de Accesibilidad Geográfica usando el modelo E2SFCA en hospitales localizados en Armenia, Quindío* (Barríos et al., 2023)

La investigación tiene como propósito evaluar las condiciones actuales y futuras de accesibilidad geográfica a servicios de salud en la ciudad de Armenia, ubicada en el departamento del Quindío, Colombia. Para realizar el estudio, se emplearon datos provenientes de diversas fuentes oficiales: información demográfica del DANE, datos hospitalarios del registro especial de prestadores de servicios de salud, y datos geográficos sobre la red vial proporcionados por la Secretaría de planeación del Quindío. La zona de estudio cuenta con una población aproximada de 295,208 habitantes distribuidos en bloques urbanos.

La contribución principal radica en aplicar el método del área de captación flotante en dos pasos mejorado (E2SFCA). Este modelo avanzado permite analizar de manera detallada la accesibilidad considerando variables como la distribución espacial de hospitales, capacidad hospitalaria, características de la red vial y distribución poblacional según bloques urbanos y estratos socioeconómicos. El modelo E2SFCA se implementó en tres etapas principales: validación y actualización de datos, análisis geoespacial generando isócronas de accesibilidad en intervalos de 5, 10 y 15 minutos mediante, utilizado ArcGIS®; y finalmente el cálculo detallado del índice de accesibilidad empleando matrices origen-destino y funciones gaussianas para evaluar cómo se degrada la accesibilidad según aumenta el tiempo de viaje.

Los resultados revelaron que, en el escenario actual, Armenia presenta una accesibilidad médica insuficiente con apenas. Al implementar un escenario futuro con dos hospitales adicionales estratégicamente ubicados, se logró aumentar la accesibilidad, beneficiando especialmente a los estratos socioeconómicos más vulnerables y reduciendo las desigualdades de acceso en la ciudad.

- *Evaluación cuantitativa de la accesibilidad espacial de diversos servicios médicos urbanos basada en Big Data de citas médicas ambulatorias* (Sui et al., 2023)

El objetivo de este trabajo es evaluar la accesibilidad espacial a los servicios médicos en la ciudad de Xiamen, China. Para ello, se utilizaron datos demográficos procedentes de diversas fuentes oficiales, como la Oficina municipal de recursos naturales y planificación de Xiamen, la Plataforma unificada de citas médicas ambulatorias; y la API de planificación de rutas de Baidu Map (lbsyun.baidu.com), que permite calcular tiempos de viaje en condiciones reales del tráfico. La zona estudiada se divide en seis distritos administrativos, que comprenden un total de 504 comunidades.

La principal contribución del estudio radica en la mejora del método del área de captación flotante en dos pasos (2SFCA, por sus siglas en inglés), mediante la inclusión de variables detalladas como la estructura etaria de la población, tipos específicos de servicios médicos (medicina interna, cirugía, ginecología y obstetricia, pediatría), grandes volúmenes de datos sobre citas médicas ambulatorias y la consideración de diferentes tiempos de viaje, según el nivel hospitalario.

Los resultados mostraron que con el modelo 2SFCA tradicional aproximadamente la mitad de las comunidades presentan una buena accesibilidad general, concentrándose principalmente en la isla principal de Xiamen. Sin embargo, al utilizar

el método refinado se revelaron patrones espaciales más complejos, destacando diferencias significativas en la accesibilidad dependiendo del tipo específico de servicio médico analizado. Esto sugiere que el método tradicional puede sobreestimar considerablemente la accesibilidad comparado con el método refinado. El estudio resalta la importancia del método refinado, al proporcionar una evaluación más precisa y detallada, que permite identificar áreas con necesidades específicas y contribuir de manera efectiva al desarrollo equitativo de las ciudades en relación al acceso a los servicios médicos.

- *Acceso espacial a los servicios de Cirugía General de Emergencia y comportamientos de bypass en California* (Han et al., 2023).

El estudio analiza el acceso espacial a los servicios de Cirugía General de Emergencia (EGS) en California, EE.UU., y el comportamiento de los pacientes que optan por acudir a hospitales más lejanos en lugar de los más cercanos (*bypass behaviours*).

Para evaluar estos patrones, se empleó un modelo jerárquico espacial bayesiano (BYM2), que permite analizar la relación entre los comportamientos de bypass y diversos factores a nivel individual y geográfico. El acceso espacial a los servicios de emergencia se calculó mediante un modelo gravitacional y una variante del método de zona de captación flotante en dos pasos (E2SFCA), considerando variables como el tiempo de viaje, la capacidad hospitalaria, la demanda poblacional, la tasa de pobreza, el tipo de seguro médico, la clasificación rural-urbana, el volumen quirúrgico del hospital y la complejidad de la enfermedad. El estudio se basó en datos de hospitales, registros de admisiones y variables socioeconómicas, obtenidos de diversas fuentes oficiales.

Los resultados revelan que un mayor acceso espacial a hospitales de emergencia reduce la probabilidad de bypass, mientras que los pacientes en zonas rurales, con seguro privado o con enfermedades más complejas son más propensos a evitar los hospitales cercanos. Además, se observó una clara agrupación geográfica de estos comportamientos, lo que sugiere que una mejor distribución de los recursos hospitalarios podría ayudar a reducir la desigualdad en el acceso a la atención quirúrgica de emergencia.

- *Vulnerabilidades sociales y acceso espacial a la atención primaria de salud a través del sistema de transporte público y automóvil en el área metropolitana de Albuquerque, NM: Evaluación de disparidades mediante SIG y modelado multinivel* (Lardier et al., 2023).

Esta investigación analiza cómo las vulnerabilidades sociales afectan el acceso a la atención primaria de salud en Albuquerque, Nuevo México. En particular, examina la accesibilidad entre quienes utilizan automóvil y quienes dependen del transporte público, con el fin de identificar desigualdades geográficas y socioeconómicas en la distribución de los servicios de salud; los cuales son esenciales para la prevención, diagnóstico y tratamiento de enfermedades.

El estudio incluye los condados de Bernalillo, Sandoval y Valencia, una región con aproximadamente 1 millón de habitantes. La investigación se basa en datos de 293 ubicaciones atención primaria de salud con 1,166 médicos, junto con información

sociodemográfica y de transporte. Los datos provienen del Identificador Nacional de Proveedores, la Encuesta de la Comunidad Estadounidense, el Índice de Vulnerabilidad Social del CDC, y la red de transporte público de Albuquerque.

El estudio aplicó un modelo de Área de Captación Flotante Mejorado en Dos Pasos (MM-E2SFCA por sus siglas en inglés) para medir la accesibilidad considerando el tiempo de viaje en automóvil y autobús. Además, se emplea un Modelo Multinivel Generalizado (GLMEM por sus siglas en inglés) para evaluar cómo las condiciones socioeconómicas afectan el acceso a los servicios médicos. Se consideraron variables como la proporción de grupos étnicos, el porcentaje de hogares liderados por mujeres, el nivel educativo y el acceso a infraestructura de transporte. Se utilizaron los Sistemas de Información Geográfica para mapear la accesibilidad espacial y un modelo multinivel para analizar la relación entre vulnerabilidad social y acceso a APS.

Los resultados muestran que el acceso es significativamente menor para quienes dependen del transporte público en comparación con quienes tienen automóvil. El estudio concluye que las desigualdades en el acceso están influenciadas tanto por factores geográficos como socioeconómicos.

La Tabla 2.2, Tabla 2.3 y la Tabla 2.4 presentan un resumen de la revisión bibliográfica realizada, enfocándose en los modelos y métodos implementados para evaluar la accesibilidad a los servicios.

Tabla 2.2. Resumen de los modelos y métodos de accesibilidad, parte 1

Trabajo	Modelos y método de accesibilidad	Variables	Contexto de la accesibilidad
(Xia et al., 2018)	Modelo integrado de Gravedad-Radiación	Costos de viaje, población y red vial	Ciudades
(Biosca Espinosa, 2018) &	Indicador de separación Indicador de interacción espacial	Tiempo de viaje, distancias, costos de transporte, modalidad de transporte	Entornos periféricos
(Logan et al., 2019)	Algoritmo de enrutamiento de código abierto	Modalidad de transporte, distancias y tiempos de viaje, información demográfica	Proximidad a servicios: escuelas secundarias públicas, atención médica y parques
(Dejen et al., 2019)	Escenarios de viaje: umbrales de tiempo ArcGis 9.3.1 ERDAS IMAGINE 9.1.	Cobertura del suelo, el modelo de elevación digital, la pendiente de la calle, la distribución de la población, el acceso por la red vial, la ubicación geográfica de los servicios de salud y el escenario de viaje	Centros de servicios de salud
(Quezada et al., 2019)	Método de interpolación Inverse Distance Weighting (IDW)	Tiempo, distancia, velocidad, número de equipamientos	Equipamientos: las clínicas de salud, hospitales, centro de salud, establecimientos públicos y privados de educación, mercados, farmacias, perfumerías, etc.
(Basso et al., 2020)	Indicador de accesibilidad: número de oportunidades que se pueden alcanzar en un tiempo de viaje determinado	Tiempos de viaje, número de oportunidades alcanzables	Educación, la salud y las oportunidades laborales

Tabla 2.3. Resumen de los modelos y métodos de accesibilidad, parte 2

Trabajo	Modelos y método de accesibilidad	Variables	Contexto de la accesibilidad
(Kaplan et al., 2020)	Medida de centralidad de cercanía basada en gráficos Análisis de Componentes Principales Factorización del eje principal	Modalidad de viaje, caminos pavimentados,	Accesibilidad espacial de la vía nacional multiescalar
(Montes et al., 2020)	Modelo p-mediana con restricciones de distancia Sistemas de Información Geográfica	Distancia y variabilidad en las distancias, oferta educativa, demanda poblacional, red vial, costos de transporte y distancias óptimas	Establecimientos educativos de educación primaria
(Tao et al., 2020).	Método de área de captación flotante de dos pasos (2SFCA) Modelo jerárquico de accesibilidad (H2SFCA)	Capacidad de las instalaciones, la demanda de la población, la distancia y los niveles de atención médica	Atención médica
(Rodríguez-Segura & Kim, 2021)	Algoritmo de código abierto	Ubicación geográfica de las escuelas, distancias y la distribución geográfica de la población	Desiertos educativos
(Yenisetty & Bahadure, 2021)	Distancia	Distancias, transporte público	Hospitales, escuelas, parques, estación de policía, oficina de correos, salones comunitarios, la infraestructura vial, entre otros
(Ouma et al., 2021)	Análisis de costo-distancia	Carreteras, cuerpos de agua, ríos, topografía, población, modalidad de transporte	Atención médica
(Wang et al., 2021).	Índice estadístico Modelo de proximidad espacial Modelo de interacción espacial	Ubicación de los parques, tamaño, densidad poblacional, distancias de viaje a pie y en vehículo, patrones de urbanización, la red vial y los umbrales de distancia	Parques urbanos
(Gaxiola et al., 2021).	Urban Mobility Accessibility Computer	Distancias e indicadores socioeconómicos	Escuelas, mercados, hospitales, etc.
(Bruno et al., 2022)	Indicadores espaciales Sistemas de Información Geográfica	Ubicación de las escuelas, la distribución de la población, la infraestructura vial, los ingresos per cápita y la cobertura de internet.	Educación secundaria superior
(Mahuve & Tarimo, 2022)	Modelo FFCA	Ubicación de los puntos de servicio, la distribución poblacional y la red vial	Puntos de servicios: hospitales, comercios, oficinas gubernamentales, servicios médicos y comerciales
(Wu et al., 2023)	Modelo de barrera espacial ponderado Sistemas de Información Geográfica	Red vial, la jerarquía de carreteras, los tiempos de viaje reales en cada estación, la ubicación de los hospitales, la densidad poblacional y la eficiencia del transporte en condiciones climáticas adversas	Servicios de salud en condiciones climáticas
(Marwal & Silva, 2023)	Modelo E-2SFCA	Distancia, la riqueza, la proporción de población de castas registradas, densidad de población, ubicación espacial de los vecindarios y otras características socioeconómicas.	Servicios urbanos esenciales: escuelas, hospitales, instalaciones de entretenimiento y centros de empleo

Tabla 2.4. Resumen de los modelos y métodos de accesibilidad, parte 3

Trabajo	Modelos y método de accesibilidad	Variables	Contexto de la accesibilidad
(Shao & Luo, 2023)	Modelo E2SVCA	Número de médicos disponibles, cantidad de habitantes por bloque urbano, tiempo estimado de viaje hacia los centros de salud, los niveles de complejidad hospitalaria y la estratificación socioeconómica	Telesalud
(Barrios et al., 2023)	Modelo E2SFCA	Distribución espacial de hospitales, capacidad hospitalaria, características de la red vial y distribución poblacional según bloques urbanos y estratos socioeconómicos	Servicios de salud
(Sui et al., 2023)	método refinado de zona de captación flotante en dos pasos (E2SFCA)	Edad poblacional, diferentes tipos de servicios médicos, citas ambulatorias, distintos tiempos de viaje , nivel del hospital.	Servicios médicos
(Han et al., 2023)	Modelo jerárquico espacial bayesiano (BYM2) y método de zona de captación flotante en dos pasos (E2SFCA)	Tiempo de viaje, la capacidad hospitalaria, la demanda poblacional, pobreza, la clasificación rural-urbana, complejidad de la enfermedad	Hospitales con servicios de emergencia.
(Lardier et al., 2023)	Modelo de área de Captación Flotante Mejorado en Dos Pasos (MM-E2SFCA) Modelo Multinivel Generalizado (GLMEM) Sistemas de Información Geográfica	Proporción de grupos étnicos, el porcentaje de hogares liderados por mujeres, el nivel educativo y el acceso a infraestructura de transporte, modalidad de transporte	Atención primaria de salud

2.4. Modelos de accesibilidad a servicios mediante la Toma de Decisiones Multicriterio

- *Un enfoque multicriterio para evaluar la accesibilidad del transporte público a nivel estratégico* (Hawas et al., 2016).

Este trabajo tiene como objetivo evaluar la accesibilidad al transporte público en diversos distritos de la isla de Abu Dhabi. La investigación se centra específicamente en cómo la ubicación de las paradas de autobús, la oferta de autobuses, la cobertura de los servicios y la diversidad de rutas satisfacen las necesidades de la población en sus desplazamientos diarios.

Para el análisis, se consideraron tres variables principales: cobertura, oferta y diversidad de rutas, cada uno de los cuales incluye varias variables. La zona de estudio se dividió en áreas con características similares en cuanto a transporte público. En total, se analizaron 33 distritos con 47 rutas de autobús. Los datos necesarios fueron proporcionados por el Departamento de Transporte de Abu Dhabi y el Consejo de Planificación Urbana de Abu Dhabi. Para la estimación de diversos indicadores espaciales, se utilizó el Sistema de Información Geográfica ArcGis.

Dentro de la metodología, se llevó a cabo una consulta con profesionales y expertos en transporte público para asignar un peso relativo a los indicadores de cada variable. En el análisis de la ponderación, se utilizó el método TOPSIS, tanto para la evaluación individual como conjunta de las tres variables. Posteriormente, se aplicó el análisis de conglomerados K-means para identificar grupos homogéneos basados en las puntuaciones que arrojó el método TOPSIS. Los resultados permitieron clasificar cada distrito en cinco categorías de accesibilidad: muy buena, buena, moderada, mala y muy mala. Finalmente, se realizó una representación gráfica para facilitar la visualización de los resultados.

- *Análisis de decisión multicriterio basado en SIG para analizar la accesibilidad de los discapacitados en la Gran Área del municipio de Irbid, Jordania* (Alzoubi et al., 2019).

El propósito de esta investigación es integrar un Sistema de Información Geográfica (por sus siglas en inglés SIG) y el Proceso de Análisis Jerárquico para modelar la accesibilidad de las personas con discapacidad en el área de Irbid, Jordania; asimismo, proporcionar una herramienta de apoyo a la toma de decisiones de los planificadores.

La recopilación de datos provino de diversas fuentes, incluyendo bases de datos geográficas de Irbid, modelos de elevación digital, asociaciones de discapacitados, centros comerciales y áreas circundantes. De acuerdo con la revisión de la literatura, se identificaron los principales servicios públicos necesarios para personas con discapacidad, tales como transporte, salud, educación, compras, recreación, deporte, instituciones públicas, asociaciones de discapacitados y edificios religiosos. Además, los investigadores definieron varios factores potenciales relacionados con la accesibilidad en el entorno físico. Se realizó un análisis de accesibilidad basado en factores espaciales, específicamente la distancia a pie a los servicios, el ancho de las aceras y la pendiente.

Posteriormente, se aplicó la técnica de Análisis de Decisión de Criterios Múltiples basada en SIG para generar el mapa final de accesibilidad. Este proceso incluyó la creación de una matriz de comparaciones utilizando la escala de nueve niveles de Saaty y el cálculo del índice de consistencia. A partir de esto, se representó la accesibilidad a través niveles en un mapa que muestran los diversos servicios y la accesibilidad general. El análisis de los datos se realizó utilizando ArcGis 10.3.

- *Evaluación de las Desigualdades en la Calidad de Vida: Un Enfoque Geográfico* (Faka, 2020).

Este artículo se enfoca en analizar las desigualdades geográficas en la calidad de vida, mediante la integración de factores socioeconómicos, ambientales, demográficos y de infraestructura en diversas regiones. Las variables que consideró son acceso a servicios, calidad ambiental, empleo, educación, salud y distribución de infraestructura. Estas variables fueron analizadas mediante herramientas de análisis espacial, estadísticas descriptivas y modelos geográficos.

El enfoque utilizado fue el análisis de decisiones multicriterio y el proceso de jerarquías analíticas. Los métodos empleados fueron los Sistemas de Información Geográfica (SIG) para la representación y análisis de datos georreferenciados. También se

emplearon técnicas de análisis de patrones espaciales para identificar concentraciones de desigualdad y métricas multidimensionales que permiten comparar regiones de manera uniforme. Asimismo, se utilizaron modelos de correlación espacial y análisis de densidad para comprender cómo se relacionan las variables en el espacio.

Este trabajo ofrece una metodología práctica para contribuir al diseño de políticas públicas más equitativas. Resalta la importancia de un enfoque territorial en la toma de decisiones, enfatizando que los contextos locales son esenciales para implementar estrategias que reduzcan desigualdades y fomenten un desarrollo sostenible e inclusivo.

- *Un enfoque de toma de decisiones basado en múltiples criterios para cuantificar y clasificar la accesibilidad a los servicios de mercado en zonas rurales: una aplicación en Bangladesh* (Zafri et al., 2021).

Este trabajo se centra en cuantificar y clasificar la accesibilidad a las instalaciones de mercado en áreas rurales. El enfoque comienza con la identificación de los factores que influyen en la accesibilidad a estas instalaciones. Posteriormente, se aplica el método de Comparación por Pares de Suma Constante (CSPCM) para determinar el peso relativo de cada factor. Con base en estos pesos, se utiliza la técnica de Preferencia de Orden de Similitud con una Solución Ideal (TOPSIS) para evaluar y clasificar la accesibilidad a las instalaciones. El caso de estudio se realiza en Bangladesh.

Para evaluar la accesibilidad, se seleccionaron 10 subdistritos dispersos en Bangladesh. Los mercados se categorizaron en dos tipos: mercado local y mercado de acopio. La selección de los factores que influyen en la accesibilidad se realizó a través de dos debates por cada subdistrito, considerando personas con diversos perfiles socioeconómicos. Además, se llevaron a cabo encuestas a informantes y se consultó a expertos para obtener opiniones calificadas. La comparación por pares realizada con expertos permitió definir los factores finales para el análisis de accesibilidad: distancia de viaje, tiempo de viaje, costo de viaje y estado de la carretera. Con esta información, se aplicó el método CSPCM para asignar pesos a cada factor.

Finalmente, se utilizó el método TOPSIS para cuantificar y clasificar la accesibilidad a las instalaciones de mercado. Este método se basa en seleccionar la alternativa que tenga la distancia más cercana a la solución ideal positiva y la distancia más lejana a la solución ideal negativa. Los resultados finales ofrecen un ranking de los subdistritos en términos de accesibilidad a los mercados, destacando su proximidad a estos.

- *Evaluación de la accesibilidad local de viviendas para personas mayores utilizando un enfoque multicriterio: evidencia de la República Checa* (Vaňková & Vavrek, 2021).

Esta investigación se centra en la evaluación de la disponibilidad de viviendas para personas mayores. Como caso de estudio se consideró al República Checa. Los datos recolectados provienen de fuentes del ministerio de trabajo y asuntos sociales de la República Checa, oficina de estadística y del instituto de información de la salud. A partir de la información anterior, se consideraron 14 indicadores relacionados con los

servicios de asistencia social, las condiciones sociodemográficas, necesidades geográficas de las áreas. En el estudio se consideraron 77 distritos y 419 viviendas para las personas mayores.

Para la evaluación se utilizó un enfoque, específicamente se empleó la Técnica de Orden de Preferencia por Semejanza con la Solución Ideal en combinación con el Método de Coeficiente de Varianza (CV-TOPSIS), con un enfoque objetivo. El resultado mostró el peso relativo de cada indicador, el cual permite medir su importancia en el contexto de estudio. Por tal motivo, a partir de la evaluación total de los datos, se obtuvieron los distritos con menor y mejor accesibilidad a las viviendas para personas mayores.

- *Integración de modelos de accesibilidad y localización-asignación SIG con análisis de decisiones multicriterio para evaluar la calidad de vida en la ciudad de Buraidah, KSA (Karim & Awawdeh, 2020).*

El objetivo de esta investigación fue evaluar la calidad de vida en la ciudad de Buraidah, Arabia Saudita, mediante la integración del análisis de red basado en Sistemas de Información Geográfica y el análisis de decisiones multicriterio. La metodología empleada se estructuró en varias etapas: preprocesamiento de datos, análisis de accesibilidad, análisis multicriterio y obtención del modelo de asignación utilizando herramientas SIG.

Los datos se recopilaban de diversas fuentes, como Earth Explorer, el Municipio de Al-Qassim, afiliado al Ministerio de Asuntos Municipales y Rurales, y a la Autoridad General de Estadísticas. Entre los datos de entrada se incluyeron la red de carreteras, la división territorial de la ciudad por distritos, mapas de densidad poblacional y desarrollo urbano. Para la red de carreteras, se consideraron atributos como velocidad, distancia, tiempo de espera, tiempo de viaje y dirección. Algunos de estos atributos fueron calculados utilizando ArcGis, estableciendo umbrales de tiempo de 5, 10 y 15 minutos.

Para el análisis de localización-asignación, se empleó un modelo de optimización de cobertura máxima; el cual permitió identificar deficiencias en los servicios existentes y asignar ubicaciones óptimas para los servicios sugeridos, con el objetivo de corregir estas deficiencias.

La selección de las variables que influyen en la calidad de vida se realizó mediante una revisión de la literatura y entrevistas. Para este estudio, se consideraron siete tipos de servicios, cada uno relacionado con variables específicas: servicios de salud (hospital, centro de salud y centro de ambulancia); servicios educativos (escuela primaria, escuela secundaria, preparatoria y universidad); servicios de seguridad; servicios gubernamentales; servicios religiosos; servicios deportivos; y servicios recreativos.

Para la ponderación de variables, se utilizó la técnica de Análisis Jerárquico (AHP). Una vez obtenidos los pesos de las variables; estos datos se integraron en el entorno SIG para generar los mapas correspondientes. Los resultados mostraron niveles de calidad de vida que varían desde muy bajo, bajo, moderado, alto, hasta muy alto. Además, se realizó un análisis detallado de la accesibilidad a cada uno de los servicios.

-
- *Accesibilidad de los centros de vacunación en el control del brote de COVID-19: un enfoque de Toma de Decisiones Multicriterio basado en SIG* (Alemdar et al., 2021).

El objetivo de este trabajo es determinar la asignación de nueve centros de vacunación para el COVID-19 en Estambul, Turquía; utilizando una metodología basada en decisiones científicas y estratégicas en el análisis geográfico. La selección de los criterios fue realizada por un consejo asesor compuesto por expertos y académicos. En este estudio, se consideraron ocho variables: red vial, paradas de transporte público, estaciones de emergencia, hospitales, farmacias, población, edificios, y estaciones de policía. Los datos fueron obtenidos de diversas fuentes, y los criterios se analizaron en función de las distancias; para la población, se utilizó la densidad de kernel.

Para la ponderación de variables, se empleó el método AHP. El análisis espacial de las variables se llevó a cabo en un entorno de Sistemas de Información Geográfica (SIG). La combinación de técnicas de SIG y análisis de decisiones multicriterio proporcionó un marco científico sólido para los tomadores de decisiones. Como resultado, se presenta un mapa de idoneidad dividido en 10 clases, desde las áreas más adecuadas hasta las menos adecuadas. Finalmente, se realizó un análisis de sensibilidad para evaluar los efectos de las variables en las ubicaciones de los centros de vacunación bajo diferentes escenarios.

- *Propuesta de una medida integrada basada en la accesibilidad para evaluar la equidad espacial entre diferentes clases sociales* (Azmoodeh et al., 2021).

Esta investigación se centra en emplear una medida de accesibilidad de los residentes a las instalaciones públicas en la ciudad de Teherán, Irán, utilizando la Toma de Decisiones Multicriterio. En particular, se analizó un distrito de los 22 existentes, que cuenta con 1,183 bloques residenciales. Los datos sobre la distribución espacial, datos demográficos y de vivienda fueron recopilados del Municipio de Teherán. Con esta información, se clasificaron 10 indicadores de accesibilidad en dos categorías: basados en la distancia y no basados en la distancia.

Para calcular el índice de accesibilidad y evaluar la equidad en un área urbana de los datos basados en la distancia, se utilizó la medida propuesta por Hansen; mientras que para el resto de los datos se aplicaron el coeficiente de Gini y el Índice de Moran. Posteriormente, se realizó una comparación por pares basada en las opiniones de los tomadores de decisiones. El siguiente paso se empleó el método AHP y, finalmente, se evaluaron los datos mediante una combinación lineal ponderada para obtener el índice final. Este índice se aplicó a diferentes clases sociales y los resultados fueron representados gráficamente utilizando Sistemas de Información Geográfica (SIG).

- *Un marco de análisis de decisiones multicriterio para medir el acceso equitativo a la atención sanitaria durante el COVID-19* (Roy & Kar, 2022).

La finalidad de este trabajo fue la evaluación del acceso equitativo a los servicios de salud durante la pandemia de COVID-19, aplicando un marco basado en el análisis de decisiones multicriterio (MCDA). La metodología considera diversos factores agrupados en tres categorías: vulnerabilidad social (pobreza, densidad poblacional, acceso al transporte), características del entorno construido (propiedad de vehículos,

transpirabilidad, densidad hospitalaria) y la distribución de casos de COVID-19. El estudio se realizó en el área metropolitana de Los Ángeles.

Los datos extraídos provienen de la oficina del censo, el centro para el control y la prevención de enfermedades, portal de datos abiertos del condado de Los Ángeles, departamento de transporte y la red de carreteras de Open Street Map. Las variables fueron ponderadas para construir un índice de accesibilidad equitativa que permitió identificar áreas con desigualdades en el acceso a hospitales. Para la asignación de pesos de las variables se consideró una ponderación igual y un modelo de agregación de lineal de suma ponderada.

Como resultado, el estudio destaca las áreas con menor acceso a servicios de salud, proporcionando información clave para la toma de decisiones y la asignación de recursos durante crisis sanitarias. También, subraya la importancia de abordar las disparidades estructurales en el acceso a los servicios esenciales.

La Tabla 2.5 presenta un resumen de los trabajos de accesibilidad basados en la Toma de Decisiones Multicriterio.

Tabla 2.5. Resumen modelos de accesibilidad con enfoque Multicriterio

Trabajo	Métodos multicriterio	Variables	Contexto de la accesibilidad
(Hawas et al., 2016)	Consulta con expertos, TOPSIS, k-means	cobertura, oferta y diversidad de rutas	Transporte público
(Alzoubay et al., 2019)	Proceso de Jerárquico (AHP)	Servicios públicos de interés Aspectos físicos	Para personas con discapacidad
(Faka, 2020)	Proceso de Jerárquico (AHP)	Servicios, calidad ambiental, empleo, educación, salud y distribución de infraestructura	Desigualdad en la calidad de vida
(Zafri et al., 2021)	Método de Comparación por Pares de Suma Constante (CSPCM), Preferencia de Orden de Similitud con una Solución Ideal (TOPSIS), Encuestas	Distancia de viaje, tiempo de viaje, costo de viaje y estado de la carretera	Mercados en áreas rurales
(Vaňková & Vavrek, 2021)	Técnica de Orden de Preferencia por Semejanza con la Solución Ideal en combinación con el Método de Coeficiente de Varianza (CV-TOPSIS)	Servicios de asistencia social, condiciones sociodemográficas y necesidades geográficas	Disponibilidad de viviendas para personas mayores
(Karim & Awawdeh, 2020)	Entrevistas, Proceso de Jerárquico (AHP)	Red de carreteras, distancia, tiempo, servicios públicos	Evaluación de la calidad de vida con base a los servicios
(Alemdar et al., 2021).	Consenso con expertos, Densidad de kernel, Proceso de Análisis Jerárquico (AHP)	Transporte público, población, distancias a diversas instalaciones públicas	Centros de vacunación
(Azmoodeh et al., 2021)	Proceso de Jerárquico (AHP)	Medida de accesibilidad relativa, coeficiente de Gini, índice de Morán	Instalaciones públicas
(Roy & Kar, 2022)	Ponderación equitativa	Vulnerabilidad social, características del entorno y la distribución de casos	Hospitales

2.5. Análisis del estado del arte

La revisión de la literatura evidencia una evolución significativa en los enfoques metodológicos para evaluar la accesibilidad. Los estudios iniciales se centraban en modelos tradicionales que priorizaban medidas básicas de proximidad geográfica. Sin embargo, los trabajos más recientes han ampliado el panorama mediante la incorporación de nuevas variables, modelos analíticos avanzados y tecnologías emergentes. A lo largo de las últimas décadas, la investigación sobre accesibilidad ha mostrado un crecimiento notable en términos de diversidad de indicadores, métodos y medidas. Aunque muchos de los métodos actuales se basan en los primeros modelos de accesibilidad, han evolucionado para incorporar aspectos más complejos. Sin embargo, gran parte de la literatura sigue predominando en el uso de enfoques espaciales que priorizan la proximidad geográfica a los servicios como principal criterio de análisis. Aunque este enfoque es útil, puede resultar insuficiente, ya que no considera otros factores críticos como las características sociales, económicas, demográficas, etc. de las áreas estudiadas.

Un componente destacado en la literatura revisada es el uso de Sistemas de Información Geográfica. Estas herramientas han sido fundamentales para el análisis de accesibilidad, facilitando la representación gráfica precisa, la obtención de los datos, la integración de datos espaciales y la elaboración de análisis más detallados. Sin embargo, se identificaron limitaciones importantes relacionadas con la disponibilidad y calidad de los datos, así como con el acceso a información integral y actualizada. La ausencia de datos confiables puede llevar a conclusiones sesgadas o incompletas, subestimando las verdaderas condiciones de accesibilidad.

Adicionalmente, se identificaron pocas investigaciones que aborden la accesibilidad en contextos mexicanos. Igualmente, se observó que son limitadas las investigaciones que incorporan enfoques más allá de los análisis espaciales tradicionales, como los métodos de Toma de Decisiones Multicriterio (MCDM). En particular, los modelos multicriterio han cobrado relevancia en el contexto de la accesibilidad, al integrar múltiples dimensiones en el análisis y ofreciendo una visión más completa y detallada. Este enfoque ha permitido considerar simultáneamente diversos factores, enriqueciendo la comprensión del fenómeno y facilitando una evaluación más integral. Estos enfoques han demostrado ser herramientas poderosas para comprender la accesibilidad desde una perspectiva más holística, al permitir la incorporación de múltiples variables en el análisis. Esto ha resultado en modelos más detallados y realistas que consideran un espectro más amplio de factores que influyen en el acceso a los servicios.

A pesar de los avances logrados, persisten desafíos y vacíos que requieren atención en este campo. Entre estos se destaca la necesidad de integrar técnicas multidimensionales y enfoques innovadores, en combinación con las capacidades avanzadas de la tecnología. Tal integración puede proporcionar una comprensión más integral y precisa del acceso a los servicios, además de guiar de manera más efectiva la toma de decisiones en áreas clave como la planificación urbana, la gestión de recursos y el desarrollo sostenible.

Capítulo 3. Marco teórico

3.1. Introducción

En este capítulo se ofrece una descripción detallada de los conceptos y fundamentos teóricos empleados en esta investigación. Se presentan los elementos, principios y modelos matemáticos que sustentan la investigación, con el propósito de proporcionar un marco conceptual sólido y contextualizado para los enfoques aplicados. Los temas abordados incluyen la Toma de Decisiones Multicriterio, las técnicas matemáticas para la identificación de variables relevantes, los Sistemas de Información Geográfica y el análisis de la Accesibilidad a Servicios.

3.2. Toma de Decisiones Multicriterio

En esta sección se presentan los fundamentos de la toma de decisiones, específicamente la Toma de Decisiones Multicriterio; así como los elementos y conceptos matemáticos que la conforman.

Fundamentos de la Toma de Decisiones

La Investigación de Operaciones es una rama de las matemáticas enfocada en la aplicación de métodos y técnicas matemáticas para resolver problemas que requieren soluciones óptimas y en la determinación del mejor curso de acción. Según (Carro, 2014), la definición de Churchman, Ackoff y Arnoff describe a la Investigación de Operaciones como la aplicación del método científico a problemas relacionados con el control de organizaciones o sistemas, llevada a cabo por grupos interdisciplinarios. Su finalidad es desarrollar soluciones que se alineen de manera óptima con los objetivos de la organización. El objetivo de la Investigación de Operaciones es proporcionar una base analítica que apoye a los responsables en la resolución efectiva de problemas y la toma de decisiones informada (Rincón, 2001).

Una característica de la Investigación de Operaciones es su enfoque sistémico; el cual se centra en comprender y abordar problemas desde una perspectiva integral, considerando las relaciones entre los componentes del sistema. Este enfoque permite una visión más completa, efectiva y contribuye al funcionamiento óptimo del sistema en general. Por consiguiente, la metodología de la Investigación de Operaciones sigue un conjunto de pasos establecidos para tratar problemas complejos y encontrar soluciones óptimas. Estos pasos incluyen la formulación del problema, la construcción de un modelo matemático que represente el problema, la resolución del modelo con técnicas apropiadas, la interpretación y validación de los resultados obtenidos. Posteriormente, se procede a la implementación de las soluciones propuestas (Moroto et al., 2012).

Teoría de decisiones

La Teoría de Decisiones es una disciplina dentro de la Investigación de Operaciones, que tiene como objetivo ayudar a identificar la mejor opción en situaciones donde las decisiones deben tomarse en medio de factores bajo riesgo, incertidumbre, inciertos, en conflicto y complejos. En este marco, el análisis de decisiones aplica un enfoque racional para elegir la alternativa más adecuada, y la calidad de la opción seleccionada

depende en gran medida de la precisión de los datos utilizados para evaluar el caso en cuestión (Eppen et al., 2000), (Taha, 2004), (Baird, 1978).

La toma de decisiones es una actividad esencial del ser humano que revela su habilidad, crecimiento personal y nivel de autonomía. Debido a su naturaleza fundamental, se ha analizado desde numerosas perspectivas, lo que ha dado lugar a una amplia gama de enfoques y problemas. Un ejemplo de ello es la Teoría de decisiones desde la perspectiva de la Investigación de Operaciones (Moreno et al., 1998).

Proceso de decisión

Dentro de la Teoría de Decisiones, el Proceso de Decisión implica seleccionar una o más alternativas de entre un conjunto de opciones disponibles, basándose en criterios específicos para lograr objetivos y metas predefinidos. Este proceso se realiza dentro de un contexto que incluye posibles estados de la naturaleza o criterios, y consta de los siguientes elementos:

- o Una o más decisiones con una serie de objetivos y metas definidos.
- o Un conjunto de posibles alternativas disponibles para los decisores.
- o Un conjunto de posibles resultados que se derivan de la implementación de las acciones.
- o Un entorno compuesto por los posibles estados de la naturaleza que afectan los objetivos de las decisiones, sobre los cuales los decisores no tienen control.
- o Una función que relaciona las acciones y los resultados con el entorno.
- o Un proceso de toma de decisiones que selecciona una o varias alternativas en función de un entorno específico y metas explícitas del grupo de decisores.
- o Un criterio que define y orienta el proceso de toma de decisiones.
- o *Una variable* representa cualquier factor o parámetro que interviene, modifica o condiciona alguno de los componentes del proceso de toma de decisiones.

En la teoría de decisiones se abordan diversos tipos de problemas, lo que permite que el proceso de toma de decisiones se adapte a diferentes contextos y factores. Por lo tanto, es esencial considerar una variedad de elementos y condiciones al desarrollar el proceso de decisión y diseñar estrategias que garanticen decisiones efectivas (Cardozo, 2011) (Prawda, 1996). Por ejemplo, los problemas bajo conflicto consideran diversos: criterios y metas que compiten entre sí. La toma de decisiones en este contexto implica gestionar compromisos entre diferentes criterios para encontrar una solución que optimice el beneficio general de manera equilibrada y minimice los conflictos entre las variables en competencia.

3.2.1. Toma de Decisiones Multicriterio

La Toma de Decisiones Multicriterio, es un área centrada en abordar problemas de la teoría de decisiones, especialmente en contextos bajo conflicto. La Toma de Decisiones Multicriterio es una disciplina orientada a la ciencia de la elección, cuyo objetivo principal es identificar y seleccionar las mejores opciones basándose en los objetivos y preferencias de los tomadores de decisiones. Este enfoque busca determinar la alternativa que mejor se ajuste a sus metas, valores y prioridades, proporcionando una base sólida para tomar decisiones informadas y efectivas. Dentro de la literatura la Toma de Decisiones Multicriterio, también se define como Toma de Decisiones Multiatributo, Análisis de Decisiones Multicriterio, Toma de Decisiones Multiobjetivo y Toma de Decisiones Multidimensional (Zardari et al., 2015).

La Toma de Decisiones Multicriterio surgió en 1971 como un enfoque para comparar y seleccionar alternativas en situaciones donde deben considerarse múltiples criterios. La Sociedad Internacional de Toma de Decisiones Multicriterio define la Toma de Decisiones Multicriterio como el estudio de métodos y técnicas que permiten integrar las variables en el proceso de decisión. La Toma de Decisiones Multicriterio tiene como objetivo principal proporcionar a los que toman las decisiones herramientas que faciliten la resolución de problemas complejos, considerando diversas variables que a menudo entran en conflicto de manera simultánea (Zardari et al., 2015).

La Toma de Decisiones Multicriterio se clasifica en dos categorías principales: Toma de Decisiones Multiatributo y Toma de Decisiones Multiobjetivo. La Toma de Decisiones Multiobjetivo analiza las variables relacionadas mediante funciones objetivo, donde las alternativas se definen en función de las restricciones impuestas sobre dichas variables. Su objetivo es encontrar una solución que optimice simultáneamente los valores de las funciones objetivo y cumpla con las restricciones establecidas.

La Toma de Decisiones Multiatributo se enfoca en comparar diversas alternativas, tomando en cuenta múltiples variables y el nivel de importancia; con el propósito de evaluar dichas alternativas para establecer un ranking que facilite la selección de la opción más adecuada para el tomador de decisiones. Este enfoque es particularmente útil en contextos donde las alternativas son finitas y las variables están definidos y medibles (Malczewski, 2017). En la mayoría de los casos, el proceso de decisión en la Toma de Decisiones Multiatributo sigue los pasos que se muestran en la Figura 3.1 (Zardari et al., 2015).

Por lo anterior, esta investigación se centra en la Toma de Decisiones Multiatributo, un término que en la literatura se reconoce comúnmente como Toma de Decisiones Multicriterio. Para los fines de esta investigación, se empleará el término Toma de Decisiones Multicriterio.



Figura 3.1. Metodología de la Toma de Decisiones Multicriterio

3.2.2. Elementos de la Toma de Decisiones Multicriterio

En esta sección se describen los elementos fundamentales en el análisis de la Toma de Decisiones Multicriterio.

Estructura jerárquica de los modelos de Toma de Decisiones Multicriterio

Para determinar las preferencias de aquellos que toman la decisión, se emplean términos como metas, objetivos, criterios y variables. Las metas indican un nivel de logro deseado a futuro. Es decir, Las metas representan los resultados específicos que los tomadores de decisiones desean alcanzar, definidos en un contexto temporal y espacial determinado. Los objetivos marcan la dirección del cambio deseado. Dichos objetivos reflejan las aspiraciones generales de los tomadores de decisiones y marcan la dirección hacia la cual quieren enfocarse. Los criterios, por su parte, son resultados específicos y medibles que guían el esfuerzo hacia la mejora. Los criterios, por su parte, actúan como estándares o reglas que permiten evaluar la aceptabilidad de las alternativas, estableciendo la base para saber qué tan bien estas satisfacen los objetivos y metas definidos Finalmente, dentro de cada criterio se encuentran las variables o atributos son características específicas que se asignan a cada alternativa, los cuales pueden ser cuantitativos o cualitativos y contribuyen a la evaluación de las alternativas según los criterios establecidos (Zarghami & Szidarovszky, 2011), (Pedrycz et al., 2010), (Malczewski & Rinner, 2015). La Figura 3.2 muestra la relación jerárquica entre las metas, objetivos, criterios y variables.



Figura 3.2. Relación entre metas, objetivos, criterios y variables

Elementos de los problemas de Toma de Decisiones Multicriterio

En los problemas de Toma de Decisiones Multicriterio existen tres principales elementos (Zarghami & Szidarovszky, 2011), (Pedrycz et al., 2010), (Malczewski & Rinner, 2015) y (Moskowitz, H.; Wright, 1979). . A continuación, se describen estos componentes, véase Figura 3.3.

- **Tomador de decisiones:** En este punto se puede considerar una sola persona responsable de la decisión, mientras que en otros casos puede ser que varios individuos u organizaciones sean los involucrados. Cuando existe diversos tomadores de decisiones, es posible que ninguna opción satisfaga a todos de manera equitativa; por tal motivo se debe tomar una decisión colectiva, que satisfaga los intereses de manera conjunta.
- **Alternativas:** Estas son las opciones disponibles para elegir. El conjunto de todas las posibles alternativas se conoce como el espacio de decisiones, el cual debe estar bien definido.
- **Criterios:** Son las variables agrupadas en dimensiones que representan las características de cada alternativa. Estas variables permiten evaluar y comparar las alternativas.

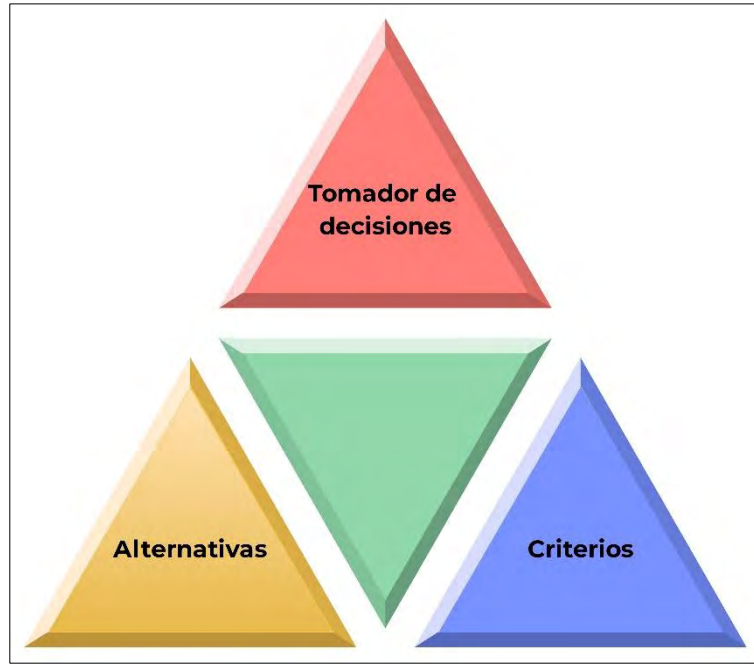


Figura 3.3. Elementos de los problemas de Toma de Decisiones Multicriterio

Componentes matemáticos para el desarrollo de un modelo multicriterio

- Matriz de decisión

Es la representación estructurada de alternativas y variables (criterios): La construcción del espacio de decisión, cuando el espacio es discreto y finito, consiste en representar las alternativas, las variables y las evaluaciones de cada alternativa con respecto a estos criterios en un matriz, conocida como matriz de decisión o matriz de evaluación. En esta matriz se organizan de manera estructurada las alternativas y los valores asociados a cada una de las variables; permitiendo una comparación directa entre las diversas opciones. En resumen, la matriz de decisión refleja las características de las diferentes alternativas de decisión, facilitando su comparación, análisis y selección (Zarghami; Szidarovszky, 2011), (Malczewski & Rinner, 2015), (Eiselt, H. A.; Sandblom, 2022). La matriz se plantea de la siguiente forma en la ecuación 3.1.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2(n-1)} & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3(n-1)} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{(m-1)1} & a_{(m-1)2} & a_{(m-1)3} & \cdots & a_{(m-1)(n-1)} & a_{(m-1)n} \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{m(n-1)} & a_{mn} \end{pmatrix} \quad (3.1)$$

Donde:

La matriz A es una matriz de dimensión $m \times n$.

a_{ij} es del desempeño de la alternativa i cuando se evalúa en la variable j (valor del costo beneficio), con $i = 1, \dots, n, j = 1, \dots, m$.

Para el caso de los modelos multicriterio con aplicaciones en espaciales, cada alternativa está definida por un conjunto de coordenadas geográficas que representan la ubicación geográfica de dicha alternativa en el espacio geográfico.

- Asignación de pesos

Este proceso consiste en asignar un peso o valor a cada una de las variables de un conjunto de datos, con el objetivo de reflejar el nivel de importancia de cada variable en relación con el resto. De este modo, se busca identificar aquellas variables que tienen una mayor influencia en el proceso de toma de decisiones. La asignación de pesos se lleva a cabo mediante la implementación de métodos de ponderación de variables, que pueden ser seleccionados según el enfoque adoptado.

Dentro de la literatura existen diversas técnicas y métodos para asignar los pesos a las variables, comúnmente conocidas como métodos de ponderación de variables. Algunos de estos métodos adoptan un enfoque subjetivo, basado en la experiencia y el conocimiento de expertos en el tema. Por otro lado, otros siguen un enfoque objetivo, basado en la información extraída directamente de los datos.

La asignación de pesos facilita la comparación de alternativas al establecer un marco sistemático que evalúa cómo cada opción contribuye al logro de los objetivos establecidos. También, permite estructurar y priorizar las alternativas según los objetivos y expectativas del decisor, brindando una base sólida para la toma de decisiones (Ishizaka & Nemery, 2013). Por lo general, los pesos de las variables deben cumplir con ciertas condiciones, tal como se especifica en la ecuación 3.2.

Sean $w_1, w_2, w_3, \dots, w_n$, los pesos asociados a cada variable

$$0 \leq w_j \leq 1 \text{ y } \sum_{j=1}^n w_j = 1 \text{ para } j = \overline{1, n} \quad (3.2)$$

Los pesos deben ser escalados de forma proporcional.

- Modelos de agregación y evaluación de alternativas

Los modelos de agregación se definen como funciones matemáticas con una estructura específica que permite equilibrar y considerar de manera sistemática tanto los pesos asignados a cada variable como los valores de las mismas; garantizando así una evaluación equilibrada y coherente de las alternativas.

Estos modelos integran los valores individuales de cada alternativa con respecto a cada variable, ponderados según los pesos asignados a cada una. Esto permite la evaluación de las alternativas y calcular un puntaje global para cada opción (Ishizaka & Nemery, 2013). A continuación, se describen algunos de los modelos de agregación más comúnmente utilizados.

Función lineal de suma ponderada (3.3)

$$f(\vec{x}) = w_1x_1 + w_2x_2 + \dots + w_{n-1}x_{n-1} + w_nx_n$$

Función de producto ponderado (3.4)

$$f(\vec{x}) = w_1x_1 * w_2x_2 * \dots * w_{n-1}x_{n-1} * w_nx_n$$

$$\begin{aligned} &\text{Función aditiva de logaritmos} \\ f(\bar{x}) &= \log(w_1x_1) + \log(w_2x_2) + \dots + \log(w_{n-1}x_{n-1}) \\ &\quad + \log(w_nx_n) \end{aligned} \tag{3.5}$$

$$\begin{aligned} &\text{Función aditiva de exponenciales} \\ f(\bar{x}) &= e^{w_1x_1} + e^{w_2x_2} + \dots + e^{w_{n-1}x_{n-1}} + e^{w_nx_n} \end{aligned} \tag{3.6}$$

Una vez que las alternativas son evaluadas y asignadas con un puntaje global, los resultados pueden organizarse en orden ascendente o descendente (ordenadas en un ranking); facilitando su comparación y posterior análisis para tomar una decisión informada. Tanto el proceso de agregación como su interpretación dependen de los pesos asignados y de la función de agregación seleccionada.

- Selección y comparación de alternativas

En el contexto de la Toma de Decisiones Multicriterio, este proceso implica evaluar cómo cada alternativa cumple con los objetivos y criterios establecidos; priorizando aquellas que mejor satisfacen las necesidades o metas del decisor. La selección de alternativas consiste en identificar la opción más adecuada, considerando diversos factores y ponderando la importancia de cada uno.

A partir de los resultados obtenidos de la evaluación de alternativas y de la evaluación global de cada una, donde las alternativas se ordenan en un ranking, el proceso de selección consiste en identificar la opción más adecuada dentro del conjunto disponible; lo que facilita una decisión informada y fundamentada.

Para mejorar y facilitar la selección y toma de decisiones en esta investigación, se implementó un método de clasificación por intervalos. Este método organiza los datos obtenidos de menor a mayor en intervalos o clases; y luego son mostrados en categorías o grupos. Este enfoque facilita la comparación y selección de alternativas de manera más clara; proporcionando una herramienta valiosa para apoyar la toma de decisiones.

En esta investigación, se utilizó el modelo de optimización de Jenks, enfocado en la clasificación por intervalos, véase la ecuación 3.7. Este modelo es reconocido por su capacidad para minimizar la varianza dentro de cada grupo y maximizarla entre grupos distintos, lo que lo convierte en una opción adecuada para analizar criterios con distribuciones y características diversas. Para resolver el modelo, se emplean comúnmente técnicas de programación dinámica (North, 2009).

$$\text{Min } E = \sum_{l=1}^r \sum_{i=1}^n (z_{il} - \bar{z}_l)^2 \tag{3.7}$$

Donde:

n : es el número total de observaciones del conjunto de datos

r : es el número de intervalos

z_{il} : es el valor de la observación i en el intervalo l

$\bar{z}_l$: es la media de los valores en el intervalo l

3.3. Técnicas para la asignación de pesos en modelos multicriterio

En esta sección se describen las técnicas empleadas para la asignación de pesos a las variables dentro de la Toma de Decisiones Multicriterio en esta investigación. Se describen los principios matemáticos y los criterios empleados para realizar una ponderación adecuada, asegurando que cada variable sea valorada de acuerdo con su relevancia en el contexto del estudio. Este enfoque garantiza un proceso de asignación de pesos sólido, fundamentado y alineado con los objetivos de la investigación.

3.3.1. Análisis de Componentes Principales

El Análisis de Componentes Principales es una técnica estadística no paramétrica propuesta por Karl Pearson en 1901, según (Hasan & Abdulazeez, 2021). Esta técnica se ha consolidado como una herramienta fundamental en la reducción de la dimensionalidad de conjuntos de datos con múltiples variables, permitiendo transformar un conjunto de n atributos o variables en un nuevo espacio de variables denominado componentes principales. Estas nuevas variables están diseñadas para ser linealmente no correlacionadas y ordenadas según la cantidad de variabilidad que capturan del conjunto de datos original, (Kherif & Latypova, 2019), (Gewers et al., 2021).

PCA se basa en la descomposición espectral de la matriz de covarianza o de correlación, dependiendo de las características y los objetivos del análisis. A través de este proceso, se obtienen los valores propios y los vectores propios. Los valores propios representan la cantidad de varianza explicada. Los vectores propios, llamados componentes principales, definen las direcciones principales de variabilidad en los datos. A partir de este proceso, la primera componente principal (primer vector propio) captura la mayor variabilidad presente en los datos originales, mientras que las componentes sucesivas explican la variabilidad restante en orden descendente (Abdelaziz et al., 2020), (Rehman et al., 2020).

Es importante mencionar que los vectores propios obtenidos en el proceso de descomposición espectral cumplen con la propiedad de ser ortonormales. Esto significa que son mutuamente perpendiculares, linealmente independientes y poseen una norma unitaria. Esta propiedad garantiza que la transformación del conjunto de datos preserve la estructura subyacente y reduzca la redundancia en la información. Además, los vectores propios mantienen su dirección bajo la transformación lineal, aunque su magnitud puede cambiar por el factor λ (valor propio). Desde una perspectiva analítica, PCA permite representar los datos en una nueva base ortonormal, donde cada componente principal es una combinación lineal de las variables originales. Esta transformación facilita la identificación de patrones y estructuras latentes en los datos. (Cartone & Postiglione, 2021), (Mahmoudi et al., 2021).

A continuación, se la formulación matemática utilizada para aplicar el Análisis de Componentes Principales (PCA).

Sea X una matriz de decisión de dimensión $m \times n$, donde cada elemento m_{ij} representa el valor de la alternativa i en la variable j .

Paso 1. Calcular de la matriz de covarianza.

$$Cov(X_i, X_j) = \frac{\sum_{k=1}^m (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j)}{m-1}$$

Donde X_i, Y_i son las variables del conjunto de datos X ; $\bar{X}_i, \bar{X}_j$ son las medias de X_i y X_j respectivamente. (3.8)
Para $i, j = 1, \dots, n$ y $k = 1, \dots, m$.

$$C = \begin{pmatrix} Var(X_1) & Cov(X_1, X_1) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Var(X_n) \end{pmatrix}$$

(Janzamin et al., 2019), (Gewers et al., 2021).

Paso 2. Descomponer la matriz de covarianzas en los valores propios y vectores propios
Considere la ecuación característica.

$$Cv = \lambda v$$

Donde v es un vector propio, λ su valor propio correspondiente y C es la matriz de covarianzas.

Los valores propios se determinan resolviendo el polinomio característico mediante la ecuación.

$$det(C - \lambda I) = 0 \quad (3.9)$$

Donde I es la matriz identidad y λ es un valor real.

Los vectores propios asociados a cada valor propio se obtienen resolviendo el sistema de ecuaciones.

$$(C - \lambda I)v = 0$$

Donde I es la matriz identidad y v es el vector propio asociado al valor propio.

(Salem & Hussein, 2019), (Aït-Sahalia & Xiu, 2019).

Paso 3. Selección de componentes principales (3.10)

Los vectores propios se ordenan según la cantidad de varianza que cada uno captura, de manera que los vectores propios con los mayores valores propios corresponden a las direcciones que explican la mayor variabilidad de los datos. Los valores propios ordenados se definen como $(\lambda_1, \lambda_2, \dots, \lambda_n)$.

La matriz de vectores propios V obtenida, definida como matriz de componentes principales; es una matriz de dimensión $n \times n$ y está compuesta por vectores linealmente independientes, cada uno con magnitud unitaria y ortogonales entre sí.

$$V = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}$$

Paso 4. Proyección de los datos

Posteriormente se aplica el Modelo General de Transformación Lineal de Datos

$$Z = XW$$

Donde

Z matriz del conjunto de datos proyectada

X matriz del conjunto de datos original

V matriz de coeficientes conformada por los vectores propios (3.11)

El resultado es una matriz Z que contiene los datos proyectados en el nuevo espacio definido por los componentes principales. Cada elemento de dimensión $m \times n$, corresponde a la representación de una observación proyectada sobre un componente principal específico.

El Análisis de Componentes Principales se utiliza principalmente para la reducción de la dimensionalidad de las variables. Sin embargo, debido a su enfoque matemático, también ha sido empleado para la identificación de variables relevantes y la ponderación de variables en el desarrollo de modelos multicriterio. Esta información es obtenida del paso 2 y 3.

3.3.2. Método de Entropía

El Método de Entropía fue desarrollado por Zeleny en 1982, y está basado en la teoría de la información de Shannon. La Entropía es una herramienta ampliamente utilizada en el análisis de decisiones multicriterio, enfocada a la obtención de los pesos de las variables, desde una perspectiva objetiva. Su principal función es cuantificar la variabilidad de las variables en un conjunto de datos, midiendo la incertidumbre o dispersión de la información para asignar pesos de manera automática y sin sesgos. A diferencia de los métodos subjetivos, la entropía se fundamenta en la variabilidad en

los datos, asegurando una ponderación imparcial y basada en la diversidad inherente de las variables (Del & Tabrizi, 2020), (Wang et al., 2022), (Rong et al., 2020).

El proceso de asignación de pesos mediante entropía implica la transformación de la matriz de decisiones y la aplicación de operaciones algebraicas que extraen el grado de variabilidad de cada variable. Al otorgar mayor peso a las variables que contienen más información relevante, este método minimiza el impacto del ruido y mejora la precisión del modelo. Como resultado, se optimiza la capacidad de generalización y robustez del análisis, proporcionando una base equitativa para la comparación de variables y favoreciendo la construcción de modelos más sólidos y eficaces (Chen, 2021), (Salehi & Izadikhah, 2014). A continuación, se describe el Método de la Entropía.

Sea A una matriz de decisión de dimensión $m \times n$, donde cada elemento a_{ij} representa el valor de la alternativa i en la variable j

(3.12)

Paso 1. Normalizar la matriz de decisión

$$n_{ij} = \frac{a_{ij}}{\sum_{i=1}^m a_{ij}}, \quad i = 1, 2, \dots, m$$

Paso 2. Calcular la entropía para cada variable

$$e_j = -K \sum_{i=1}^m n_{ij} \ln(n_{ij}), \quad j = 1, 2, \dots, n$$

(3.13)

$$K = -\frac{1}{\ln(m)} \text{ constante de entropía}$$

Paso 3: Calcular el factor de diversificación para cada variable

(3.14)

$$d_j = 1 - e_j$$

Paso 4: Calcular el vector de pesos normalizados

(3.15)

$$W_j = \frac{d_j}{\sum_{j=1}^n d_j}, \quad j = 1, 2, \dots, n$$

3.4. Técnicas de identificación de variables relevantes

La identificación de variables relevantes tiene el propósito de realizar una evaluación minuciosa para identificar y retener las variables que tienen un impacto significativo en el objetivo de la investigación. El proceso implica la aplicación de técnicas que ayuden a determinar las variables de mayor impacto. A continuación, se describen las técnicas empleadas para la identificación de variables relevantes.

3.4.1. Factor de Inflación de la Varianza

El Factor de Inflación de la Varianza es una herramienta utilizada para identificar la existencia de multicolinealidad entre las variables explicativas en un modelo de regresión múltiple. La multicolinealidad surge cuando una variable independiente tiene una alta correlación con otra u otras variables, o puede ser representada como una combinación lineal de algunas de ellas (García et al., 2006).

El VIF evalúa cuanta varianza de un coeficiente de regresión aumenta debido a la colinealidad con otras variables. Es decir, indica cuánto contribuye la correlación de una variable con otras variables a la varianza de su coeficiente estimado. Si las variables explicativas no presentan redundancia, el VIF será igual a 1. Valores entre 5 y 10 indican la posible presencia de multicolinealidad. Un VIF elevado mayor a 10 sugiere una fuerte correlación lineal entre una variable independiente y las demás en el modelo; lo que puede afectar negativamente la precisión de las estimaciones de los coeficientes de regresión (Mandeville, 2008), (Akinwande et al., 2015).

Para determinar la expresión matemática del Factor de Inflación de la Varianza, se considera R_i^2 para representar la proporción de varianza en la i -ésimo variable que está asociada a otras variables del modelo. La tolerancia del i -ésimo variable se define como 1 menos la proporción de varianza que comparte con otra variable independiente en el análisis; es decir, $(1 - R_i^2)$. Esto indica la proporción de varianza que no está explicada por las demás variables independientes. Es así como el VIF se define como el inverso de la tolerancia, proporcionando una medida de cuánto se incrementa la varianza de un coeficiente de regresión debido a la colinealidad con otras variables (O'Brien, 2007). La ecuación 3.16 muestra la fórmula para determinar el VIF.

$$VIF(x_i) = \frac{1}{(1-R_i^2)} \quad (3.16)$$

Donde:

R_i^2 es el coeficiente de determinación de la regresión de la variable x_i con respecto del resto de las variables.

3.4.2. Correlación lineal

La correlación lineal es una herramienta estadística que evalúa la relación lineal entre dos variables, en términos de la fuerza y dirección. En otras palabras, el coeficiente de correlación mide la intensidad de la relación lineal entre dos variables (Pulido & De la Vara, 2008). La ecuación 3.17 muestra el coeficiente el coeficiente de correlación de Spearman, principalmente para datos que no siguen una distribución normal. La ecuación 3.18 muestra el coeficiente de correlación de Kendall, principalmente para variables ordinales y los datos no sigue una distribución normal

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (3.17)$$

Donde:

ρ : coeficiente de correlación de spearman

d_i : entre los rangos de cada par de observaciones

n : número total de observaciones

$$\tau = \frac{(n_c - n_d)}{\frac{n(n-1)}{2}} \quad (3.18)$$

Donde:

n_c : número de pares concordantes

n_d : número de pares discordantes

n : número total de observaciones

3.4.3. Análisis de Componentes Principales

El Análisis de Componentes Principales se utiliza principalmente para la reducción de la dimensionalidad de las variables. Sin embargo, debido a su enfoque matemático, también ha sido empleado para la identificación de variables relevantes. Su descripción y fundamento matemático se presentan en la sección 3.2.3.

3.4.4. Análisis Factorial

El análisis factorial es una técnica estadística multivariada crucial para explorar las complejas relaciones entre un conjunto de variables observadas. Su objetivo principal es identificar factores latentes subyacentes que expliquen las correlaciones entre estas variables y observables. Estos factores latentes son variables no observadas que pueden interpretarse como constructos teóricos que influyen en las variables observadas (Cudeck, 2000). Uno de los propósitos clave del análisis es agrupar las variables en conjuntos que sean distintos de otros grupos y homogéneos dentro de cada grupo. En resumen, el análisis factorial busca reducir la dimensionalidad del conjunto de datos al identificar un conjunto más reducido de factores que resuman la variabilidad presente en las variables observadas. La ecuación 3.19 muestra el modelo general para poder aplicar el Análisis Factorial.

$$X = LF + E \quad (3.19)$$

Donde:

X : es la matriz del conjunto de datos original

L : es la matriz de las cargas factoriales

F : es la matriz de factores

E : es la matriz de errores

En este contexto, el análisis factorial se utiliza para determinar las variables más relevantes de un conjunto de datos; asimismo obtener los factores latentes que están relacionados a cada una de las variables.

3.5. Sistemas de información geográfica

Los datos geoespaciales son información que describe objetos o eventos localizados en la superficie terrestre, caracterizados por su posición específica. Estos datos se representan mediante la combinación de coordenadas geográficas y atributos específicos que detallan las características del objeto o evento. La ubicación de estos datos puede ser estática, como un edificio, o temporal, como el recorrido de un vehículo en movimiento (IBM, 2024). Los datos geoespaciales abarcan una amplia variedad de información sobre el mundo físico y natural, lo que los hace relevantes en

numerosos campos de aplicación, tales como el urbanismo, la demografía, la salud, el medio ambiente, y la topografía, entre otros.

Los Sistemas de Información Geográfica son herramientas informáticas que permiten almacenar, consultar, manipular, analizar y visualizar grandes bases de datos espaciales. Estas plataformas integran funciones avanzadas que van desde la cartografía automática hasta los sistemas de apoyo a la toma de decisiones (Sarría, 2024).

Los Sistemas de Información Geográfica ofrecen la capacidad de identificar patrones y relaciones espaciales en los fenómenos de estudio. Desde una perspectiva computacional, los Sistemas de Información Geográfica combinan bases de datos robustas con técnicas de modelado físico y análisis digital de imágenes, proporcionando soluciones adaptadas a diversas necesidades. Dado el creciente uso de bases de datos espaciales en una variedad de campos, existe una demanda creciente de herramientas que permitan visualizar, analizar y consultar información geoespacial de manera eficiente. Los Sistemas de Información Geográfica, junto con las tecnologías emergentes como el big data y la inteligencia artificial, continúan evolucionando para satisfacer estas necesidades, fortaleciendo su papel en la solución de problemas globales y en la mejora de la calidad de vida a través de decisiones informadas y estratégicas (Leung et al., 2020).

QGIS, o Quantum GIS (www.qgis.org), es un software de código abierto y gratuito en el ámbito de los Sistemas de Información Geográfica. Su diseño permite ejecutar múltiples funciones, desde tareas básicas hasta procesos complejos, incluyendo la recopilación, almacenamiento, procesamiento, análisis, gestión y representación de datos espaciales. Una de sus principales fortalezas es su accesibilidad y versatilidad. Al ser una aplicación multiplataforma, puede ejecutarse en distintos sistemas operativos, lo que facilita su uso en diversos dispositivos. Asimismo, su arquitectura escalable y la posibilidad de desarrollar complementos en lenguajes como Python y C++ brindan una flexibilidad notable para ajustarse a necesidades específicas (Baghdadi et al, 2018)

Otra de las ventajas de QGIS es su compatibilidad con una amplia variedad de formatos de datos geoespaciales utilizados en la industria, así como su capacidad de integración con otros sistemas y servicios Sistemas de Información Geográfica. Esta interoperabilidad favorece la colaboración y el intercambio de información entre plataformas. En conclusión, QGIS es una herramienta sólida y completa que permite explorar, analizar y visualizar datos espaciales de manera eficiente, convirtiéndose en un recurso esencial en múltiples disciplinas. Por lo que esta investigación parte del análisis y visualización de los datos se realizó en QGIS (Kim et al., 2022).

3.6. Accesibilidad a servicios

Esta sección ofrece una descripción del análisis de la accesibilidad a servicios, el cual constituye el caso de estudio abordado en esta investigación.

3.6.1. Definición de accesibilidad a servicios

El concepto de acceso se refiere al grado de ajuste entre las necesidades de los usuarios y las capacidades del sistema (Penchansky & Thomas, 1981); lo cual implica la capacidad de dichos usuarios para desplazarse a un lugar y aprovechar las diversas oportunidades, como bienes, servicios y otros recursos esenciales (Ahuja & Tiwari, 2021). Derivado de lo anterior, surge el concepto de accesibilidad: misma que ha sido ampliamente utilizada en el ámbito académico; sin embargo, no existe una definición universal debido a su dependencia del contexto y el alcance de su implementación. Por tal motivo, diversos autores han propuesto distintas definiciones desde diferentes perspectivas de análisis. En este sentido, se presentarán las algunas de las primeras definiciones de accesibilidad comúnmente utilizadas para comprender su perspectiva de análisis:

- La accesibilidad es una medida de distribución espacial de actividades en torno a un punto, ajustada a la capacidad y el deseo de los usuarios para superar la separación espacial. Por ello se define como el potencial de oportunidades para la interacción (Hansen, 1959)
- accesibilidad integral. La accesibilidad relativa se refiere como el grado en el que dos puntos en la misma superficie están conectados. Por otro lado, la accesibilidad integral se define como el grado de interconexión de un punto dado con todos los demás puntos en la misma superficie (Ingram, 1971), (Dalvin & Martin, 1976).
- La accesibilidad se define como la conexión que existe entre la ubicación del servicio o suministro y la ubicación de los usuarios; considerando el medio de transporte disponible (Penchansky & Thomas, 1981).
- La accesibilidad se compone de dos elementos. El elemento físico se refiere a la distancia geográfica entre el usuario y el servicio, reflejando la proximidad espacial y la accesibilidad locacional del servicio. El elemento social, abarca factores como la clase social, el ingreso, la educación, valores culturales y creencias del usuario, horario, la calidad, la aceptación cultural, el precio del servicio, entre otros (Garrocho, 1988).
- La accesibilidad se refiere al grado de dificultad que enfrenta la población para acceder a los servicios, considerando diversos factores; los cuales pueden clasificarse en: geográficos, administrativos, socioeconómicos, físicos y territoriales. Por lo tanto, la accesibilidad se concibe desde un enfoque integral (López & Aguilar, 2004).

3.6.2. Componentes de la accesibilidad

Dentro de la revisión de la literatura sobre la accesibilidad a servicios, se han identificado cuatro componentes fundamentales (Geurs & van Wee, 2004), (CentroGeo, 2022), (Cortés & Rojas, 2020), (Rubinstein & Baladi, 2019). Estos componentes constituyen un marco para la construcción del conjunto de datos; integrando las variables adecuadas y permitiendo una estructuración más precisa que favorece el análisis y la comprensión del estudio de la accesibilidad. A continuación, se

presenta una descripción de los componentes de la accesibilidad. La Figura 3.4 muestra de manera esquemática las relaciones entre los componentes.

- **Uso de suelo:** El componente de uso del suelo se centra en la distribución y características de las actividades y oportunidades disponibles en cada destino; como empleos, comercios, servicios sanitarios, sociales y recreativos. También incluye la demanda proveniente de los lugares de origen, donde residen los usuarios.
- Este componente aborda la interacción entre la oferta y la demanda que puede generar competencia por recursos limitados; como vacantes en empleos, escuelas o camas de hospital. Este sistema refleja cómo la localización, calidad y cantidad de estas oportunidades impactan la accesibilidad en un espacio geográfico.
- **Transporte:** El componente de transporte abarca tanto la oferta como la demanda. La oferta incluye aspectos como la localización, las características de la infraestructura, la velocidad de viaje, el número de carriles, los horarios y los costos asociados. Por su parte, la demanda considera los desplazamientos de pasajeros y mercancías.
- Este componente describe la interacción entre la oferta y la demanda de algún modo de transporte específico, reflejándose factores como distancia, el tiempo de viaje y los costos que enfrentan los usuarios al desplazarse entre un origen y un destino.
- **Temporal:** El componente temporal analiza las restricciones relacionadas con el tiempo, incluyendo la disponibilidad de oportunidades en distintos momentos del día, la semana, el año; y el tiempo que los usuarios tienen para participar en otras actividades de distinta naturaleza. Este enfoque considera cómo las limitaciones temporales influyen en la accesibilidad y el uso de recursos en un espacio geográfico.
- **Individual:** El componente individual abarca las características de las personas, como sus necesidades, capacidades y oportunidades, que influyen en su acceso a medios de transporte y las oportunidades distribuidas espacialmente. Estas características incluyen factores como edad, ingresos, nivel educativo, condición física, y presupuesto de viaje. Además, determinan la capacidad de una persona para utilizar ciertos medios de transporte (por ejemplo, conducir un coche) y acceder a oportunidades cercanas.

3.6.3. Definición de accesibilidad para esta investigación

Con base en las definiciones previas de accesibilidad, esta investigación la conceptualiza como:

"La relación que existe entre un punto de origen y un punto de destino en el espacio geográfico, determinada por la interacción de diversas variables sociales, demográficas, económicas y espaciales."

La accesibilidad a servicios puede entenderse como un fenómeno complejo que integra múltiples dimensiones representadas a través de las distintas variables. Las variables espaciales incluyen la disponibilidad de servicios, la distancia y la modalidad de transporte. Las variables demográficas se refieren a la población que demanda dichos servicios. Por su parte, los variable sociales y económicos abarcan las características socioeconómicas de la población y de la zona de estudio.

A partir del análisis de accesibilidad, es posible desarrollar un modelo matemático multicriterio que caracterice el comportamiento del acceso a los servicios y cuantifique su nivel en cada localidad. Esto facilita la evaluación del acceso de la población a los servicios, proporcionando información clave para la toma de decisiones.

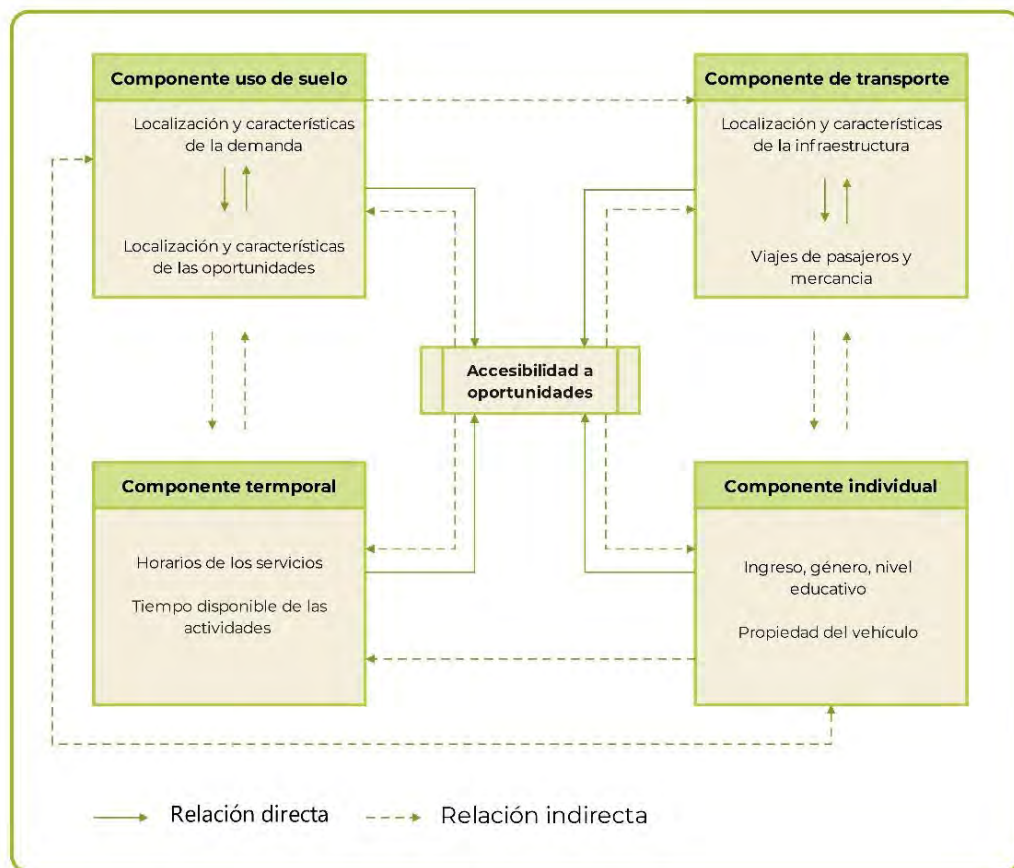


Figura 3.4. Relación entre los componentes de la accesibilidad.
Fuente: Elaborado a partir de (Geurs & van Wee, 2004), (CentroGeo, 2022), (Cortés & Rojas, 2020), (Rubinstein & Baladi, 2019).

Capítulo 4. Desarrollo del modelo matemático multicriterio

4.1. Introducción

En este capítulo se expone la parte central de la investigación, destacando dos aspectos fundamentales. El primero se enfoca en la construcción un método de ponderación de variables con enfoque objetivo para la asignación de pesos; con el propósito de determinar el nivel de relevancia de cada variable dentro del conjunto de datos. Este método, basado en un enfoque objetivo y sustentado en el análisis de datos, busca superar las limitaciones comunes de otras técnicas especialmente aquellas relacionadas con la dependencia del juicio experto; así como los desafíos asociados a

El segundo aspecto aborda el desarrollo de un modelo matemático multicriterio orientado al análisis de la accesibilidad a los servicios, sustentado en la metodología de solución planteada. Este modelo contempla la implementación del método de ponderación de variables con enfoque objetivo propuesto durante la fase de asignación de pesos. Su objetivo es integrar de manera equilibrada las variables consideradas, con el fin de representar adecuadamente el comportamiento de la accesibilidad. Una vez construido, el modelo poder llevar a cabo la cuantificación de los niveles de accesibilidad; lo que contribuye a reforzar la toma de decisiones informada en la selección de alternativas. La Figura 4.1 muestra el esquema general de la metodología de solución empleada en la investigación.

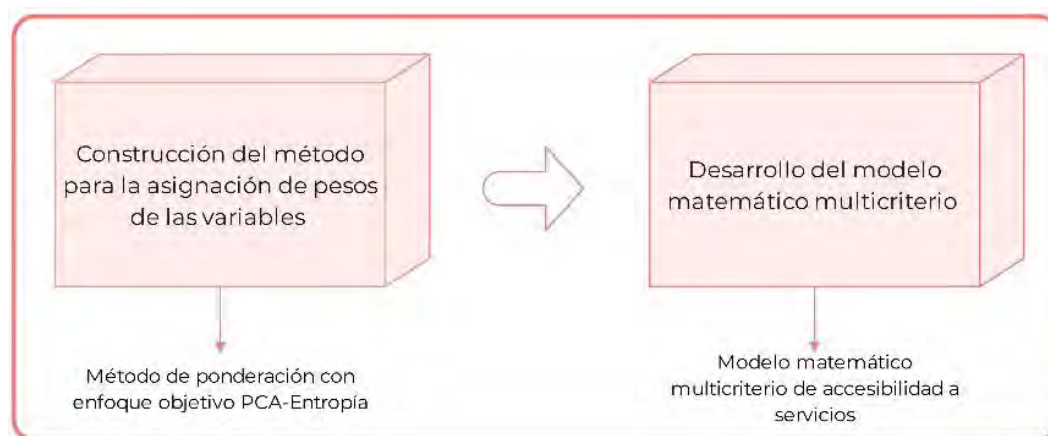


Figura 4.1. Metodología de solución de la investigación

4.2. Construcción de un método de ponderación de variables con enfoque objetivo para la asignación de pesos

En esta sección se describen los conceptos, el fundamento matemático y el desarrollo del método de ponderación de variables propuesto, el cual se basa en un enfoque objetivo para la asignación de pesos en modelos de Toma de Decisiones Multicriterio.

4.2.1. Asignación de pesos de las variables

La asignación de pesos es una etapa fundamental dentro de la Metodología de Toma de Decisiones Multicriterio, también conocida como Proceso de Toma de Decisiones

Multicriterio. La asignación de pesos implica asignar valores a las variables de un conjunto de datos, clasificándolos según su nivel de importancia. Este proceso busca evaluar el impacto relativo de cada variable en relación con las demás, dentro del conjunto de datos (matriz de decisión); con el objetivo de garantizar que la información obtenida se utilice de manera precisa en la evaluación de alternativas. Para ello, se han implementado métodos de ponderación de variables que permiten cuantificar la influencia de cada variable en el modelo; facilitando la comparación de alternativas y la selección de la opción más adecuada (Odu, 2019), (Malczewski, 2017).

La metodología de Toma de Decisiones Multicriterio (TDMC) comienza con la construcción de la matriz de decisión. El siguiente paso consiste en determinar los pesos o coeficientes que reflejan la importancia relativa de cada variable en relación con las demás, utilizando algún método de ponderación. Estos coeficientes se integran posteriormente en una función de agregación, donde se combinan con los valores de las variables asociadas a cada alternativa. Mediante esta función se calcula un valor o puntaje global para cada alternativa, lo que permite generar un ranking que ordena las opciones disponibles. Este ordenamiento facilita la identificación y selección de la(s) alternativa(s) más adecuada(s), de acuerdo con los objetivos del análisis. La Figura 4.2 presenta un esquema general que ilustra el papel que desempeña la asignación de pesos dentro de la Toma de Decisiones Multicriterio; específicamente en la formulación del modelo.

En la asignación de pesos existen diversos enfoques para calcular los coeficientes de las variables, clasificados en métodos subjetivos, objetivos e híbridos. Los métodos subjetivos están basados en la opinión de expertos. Sin embargo, sus principales limitaciones son la identificación y conformación adecuada del grupo de expertos, la posibilidad de sesgos a causa del juicio humano, entre otros. Algunos métodos de ponderación objetiva son el Proceso de Jerarquía Analítica (AHP por sus siglas en inglés), el Método Delphi, el análisis por Comparación de Pares, entre otros. Por otro lado, se encuentran los métodos objetivos. Estos métodos se basan en el análisis matemático de los datos; proporcionando una asignación de pesos fundamentada en la estructura de la información. Algunos métodos utilizados son el método de Entropía, CRITIC, Desviación estándar, Medias, MOORA, TOPSIS, ELECTRE, entre otros (Şahin, 2021), (Rădulescu & Rădulescu, 2018), (Németh et al., 2019), (Keshavarz et al., 2021).

Derivado de lo anterior, en este apartado se propone una solución innovadora para la asignación de pesos, atendiendo las limitaciones de depender de expertos. El método de ponderación de variables propuesto, llamado PCA-Entropía, combina el Análisis de Componentes Principales (PCA) con el método de Entropía; ofreciendo un enfoque híbrido y objetivo para la ponderación de atributos. Esta integración amplía las opciones existentes, proporcionando una alternativa sólida ante la falta de acceso a especialistas.

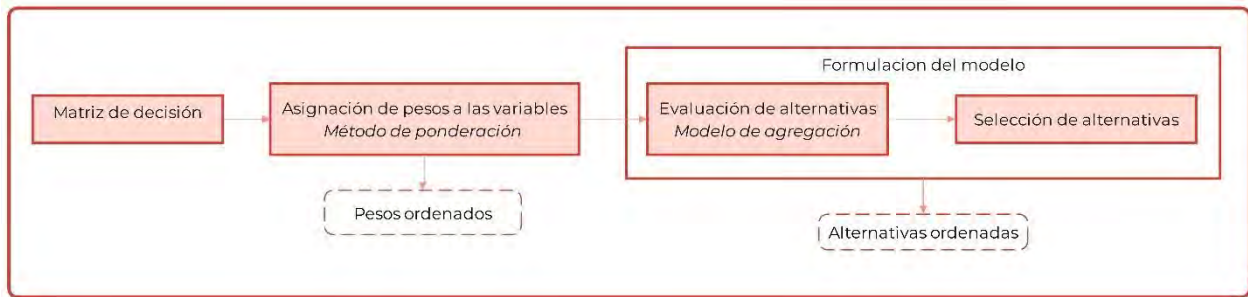


Figura 4.2. Proceso de asignación dentro de la Toma de Decisiones Multicriterio

4.2.2. Desarrollo del Método de ponderación de variables PCA-Entropía

En esta sección se ha propuesto un método para la ponderación de variables, con un enfoque objetivo. Este método considera técnicas cuantitativas que permiten asignar pesos de manera sistemática, evitando sesgos subjetivos y garantizando una representación equitativa de la información contenida en los datos. Este enfoque es especialmente útil en contextos donde no se dispone de juicio experto o se busca una justificación basada en la estructura interna de los datos.

El método propuesto, denominado como PCA-Entropía, presenta una propuesta innovadora para la ponderación de variables en conjuntos de datos con múltiples variables. Este enfoque híbrido combina las ventajas del Análisis de Componentes Principales y el Método de Entropía; ambas técnicas importantes dentro del análisis de datos.

El Análisis de Componentes Principales permite transformar un conjunto de variables correlacionadas en nuevas variables no correlacionadas, denominadas componentes principales, que revelan patrones y estructuras subyacentes en los datos. Este proceso optimiza la representación de la información, facilitando la identificación de tendencias relevantes y mejorando la comprensión general de los datos. A través de combinaciones lineales de las variables originales, se busca maximizar la varianza, mediante la obtención de los vectores propios asociados a los valores propios. De este modo, el PCA permite la reducción de la dimensionalidad sin perder información importante de las variables originales. Por otro lado, el Método de Entropía, basado en la teoría de la información, se utiliza para evaluar la cantidad de información contenida en cada variable, midiendo su dispersión o variabilidad. Al aplicar este método para ponderar variables, se minimiza el impacto del ruido en el modelo, asignando mayores pesos a aquellas variables que aportan más información.

El Método de Entropía permite una ponderación objetiva y detallada de las variables involucradas, mejorando la evaluación de su relevancia y eliminando posibles sesgos en el análisis. La combinación de este método con el PCA optimiza la representación de los datos al capturar las relaciones subyacentes entre las variables y reducir la dimensionalidad; destacando las características más significativas sin perder información crucial. A su vez, el Método de Entropía ajusta la importancia relativa de cada componente principal según la cantidad de información que aporta, asignando pesos precisos que reflejan su contribución al análisis.

Este enfoque híbrido mejora la objetividad y la precisión en la asignación de pesos, facilitando la toma de decisiones en contextos complejos. En conjunto, el método PCA-Entropía optimiza el proceso de toma de decisiones, generando modelos más robustos y confiables que pueden abordar la complejidad inherente a los entornos multicriterio.

Algunas de las razones por las que se necesita un nuevo enfoque en la ponderación de variables con métodos tradicionales:

- o *Subjetividad en la asignación de pesos:* Los métodos tradicionales de ponderación, como la asignación directa de pesos por parte de expertos o técnicas basadas únicamente en juicios humanos, presentan altos niveles de subjetividad, lo cual puede introducir sesgos y limitar la objetividad de los resultados obtenidos.}
- o *Dificultad para capturar la estructura interna de los datos:* Los métodos clásicos de ponderación no consideran las relaciones internas entre las variables, lo que puede provocar redundancia de información o sobrevaloración de ciertos criterios que presentan alta correlación entre sí.
- o *Limitaciones para manejar datos complejos y multidimensionales:* En problemas reales, los datos suelen tener múltiples dimensiones (también llamados criterios), interdependencias y patrones ocultos que los métodos tradicionales de ponderación no logran identificar ni aprovechar adecuadamente.
- o *Rigidez ante contextos dinámicos:* La estructura de los pesos tradicionales suele ser estática, dificultando su adaptación frente a escenarios donde las características de las variables o su nivel de importancia pueden cambiar con el tiempo o según el contexto del problema.

Ventajas de integrar PCA y Entropía en la ponderación de variables:

- o *Reducción de la redundancia:* El Análisis de Componentes Principales permite transformar las variables originales en nuevas componentes no correlacionadas, reduciendo la duplicación de información y facilitando una representación más eficiente de los datos.
- o *Ponderación objetiva basada en la información contenida en los datos:* El método de Entropía permite asignar pesos a las variables en función de su capacidad discriminativa, es decir, cuanto mayor variabilidad o información aporta una variable, mayor será su peso.
- o *Mejor adaptación a entornos complejos y dinámicos:* La integración del Análisis de Componentes Principales y Entropía permite capturar mejor las estructuras internas de los datos y asignar pesos de manera dinámica, reflejando tanto la importancia estadística como la capacidad informativa de cada variable.
- o *Fortalecimiento de la objetividad y neutralidad en la toma de decisiones:* La combinación PCA-Entropía elimina la influencia directa de juicios subjetivos, permitiendo que los pesos asignados respondan directamente a las propiedades intrínsecas de los datos analizados.
- o *Complementariedad metodológica:* Mientras PCA reduce la dimensión y elimina la correlación, la Entropía asigna pesos según la variabilidad de las variables restantes, logrando así una ponderación robusta, equilibrada y coherente con la naturaleza del problema.

- o *Menor susceptibilidad a las técnicas de normalización:* A diferencia de otros métodos tradicionales de ponderación, la integración de PCA-Entropía presenta menor sensibilidad a las técnicas de normalización de los datos, ya que ambos métodos se basan principalmente en las propiedades internas de la información y no dependen directamente de la escala o magnitud original de las variables.

A continuación, se describe de manera general el procedimiento del método PCA-Entropía y su desarrollo matemático. En la Figura 4.3 se muestra el esquema general del proceso del método.

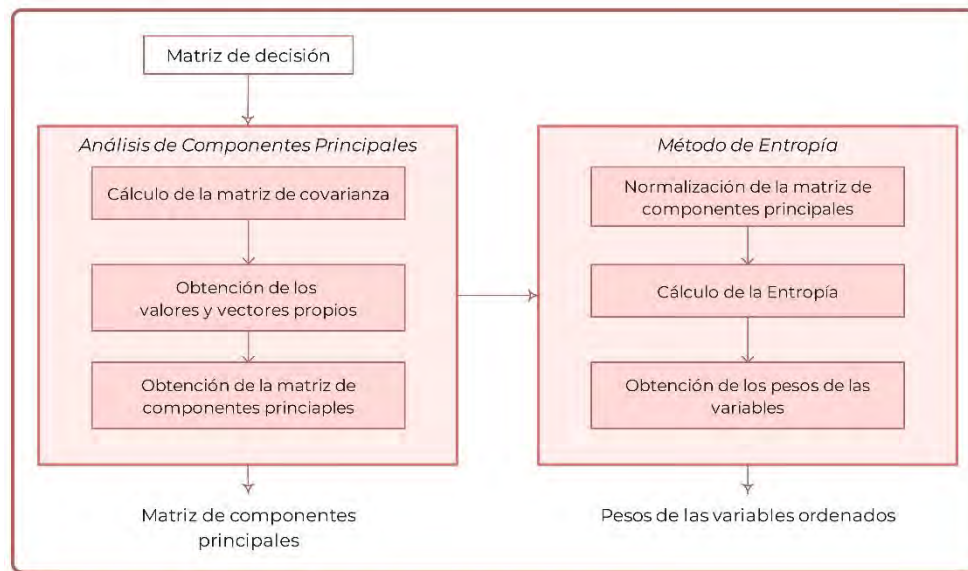


Figura 4.3. Método ponderación de variables PCA-Entropía

Descripción del método de ponderación de variables PCA-Entropía

El método se estructura en una serie de pasos secuenciales, cada uno de ellos diseñado específicamente para optimizar el proceso de evaluación y asignación de pesos a las variables, en el contexto de la Toma de Decisiones Multicriterio.

1. Cálculo de la matriz de covarianza: Se calcula la matriz de covarianzas para identificar las relaciones lineales entre las variables.
2. Obtención de los valores y vectores propios: Se descompone la matriz de covarianza a partir del cálculo de sus valores y vectores propios, lo cual permite identificar las direcciones principales de variabilidad de los datos. Los valores propios permiten determinar el porcentaje de varianza explicada por cada componente, mientras que los vectores propios definen las combinaciones lineales de las variables originales que mejor representan dicha variabilidad.
3. Obtención de la matriz de componentes principales: Se seleccionan los vectores propios más relevantes para construir esta matriz de componentes principales.
4. Normalización de la matriz de componentes principales: Se normalizan los valores para facilitar su comparación.

5. Cálculo de la entropía: Se evalúa la variabilidad de los componentes principales mediante la entropía.
6. Obtención de los pesos de las variables: Se determinan y normalizan los pesos de las variables s , reflejando su importancia relativa en el análisis multicriterio.

Desarrollo matemático del método de ponderación de criterios PCA-Entropía

Considere la matriz de datos X , que se define como la matriz de decisión y está descrita por la ecuación 4.1 con dimensión $m \times n$.

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad (4.1)$$

Donde n son las variables y m las instancias

Paso 1. Cálculo de la matriz de covarianza

A partir de la matriz de datos transformada, calcular la matriz de covarianza con base a la ecuación 4.2.

$$Cov(X_i, X_j) = \frac{\sum_{k=1}^m (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j)}{m-1} \quad (4.2)$$

Donde X_i, Y_i son criterios del conjunto de datos X ; $\bar{X}_i, \bar{X}_j$ son las medias de X_i y X_j respectivamente. Para $i, j = 1, \dots, n$ y $k = 1, \dots, m$

La matriz de covarianza C obtenida es una matriz simétrica de dimensión $n \times n$ que representa la relación de covarianza entre cada par de variables del conjunto de datos X . Cada elemento de la matriz indica el grado de variabilidad conjunta entre dos variables, proporcionando información sobre su correlación. Esta relación está expresada por la ecuación 4.3 (Janzamin et al., 2019), (Gewers et al., 2021).

$$C = \begin{pmatrix} Var(X_1) & Cov(X_1, X_1) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Var(X_n) \end{pmatrix} \quad (4.3)$$

Paso 2. Cálculo de los valores y vectores propios

Este paso consiste en descomponer la matriz de covarianza en términos de sus valores y vectores propios. Los valores propios representan la cantidad de varianza explicada por cada componente, mientras que los vectores propios definen la dirección en la que existe mayor variabilidad en los datos. El cálculo

de los valores y vectores propios se realiza utilizando la ecuación característica 4.4 (Salem & Hussein, 2019),(Aït-Sahalia & Xiu, 2019).

$$Cv = \lambda v \quad (4.4)$$

Donde v es un vector propio y λ su valor propio correspondiente.

Los valores propios se determinan resolviendo el polinomio característico mediante la ecuación 4.5.

$$\det(C - \lambda I) = 0 \quad (4.5)$$

Donde I es la matriz identidad y λ es un valor real.

Los vectores propios asociados a cada valor propio se obtienen resolviendo el sistema de ecuaciones de la 4.6.

$$(C - \lambda I)v = 0 \quad (4.6)$$

Donde I es la matriz identidad y v es el vector propio asociado al valor propio.

Los valores propios obtenidos forman una matriz diagonal de dimensión $n \times n$; donde los fuera de la diagonal principal son cero, y los elementos de la diagonal contienen los valores propios $(\lambda_1, \lambda_2, \dots, \lambda_n)$. La matriz diagonal de valores propios está descrita por la ecuación 4.7.

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad (4.7)$$

La matriz de vectores propios V obtenida, definida como matriz de componentes principales; es una matriz de dimensión $n \times n$ y está compuesta por vectores linealmente independientes, cada uno con magnitud unitaria y ortogonales entre sí. Cada columna v_i de la matriz V es un vector propio correspondiente al valor propio λ_i , véase la ecuación 4.8.

$$V = \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{pmatrix} \quad (4.8)$$

Paso 3. Obtención de la matriz de componentes principales

Este paso consiste en la obtención de la matriz de componente principales D . Para ello, se seleccionan los r primeros vectores propios de la matriz de vectores propios V , tomando como criterio el porcentaje de varianza explicada. El porcentaje de varianza explicada está determinado por los valores propios λ_i . La fracción de la varianza explicada por las r ($r < n$) primeras componentes principales están dada por la ecuación 4.9.

$$R^2 = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (4.9)$$

Donde λ_i representa el i -ésimo valor propio de la matriz de covarianza.

La matriz de componente principales D tiene dimensión de $n \times r$ y se define con base en la ecuación 4.10.

$$D = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1r} \\ v_{21} & v_{22} & \cdots & v_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nr} \end{pmatrix} \quad (4.10)$$

Donde cada columna v_i representa un vector propio correspondiente a una de las r primeras componentes principales, y cada fila representa la contribución de una variable original en estas componentes.

Paso 4. Normalización de la matriz de componente principales

En este paso, se transforman los datos de la matriz de componente principales D para garantizar que todos los valores sean positivos y se aplica una técnica de normalización basada en la suma por columna. Este procedimiento estandariza los valores, asegurando su coherencia y permitiendo una comparación equitativa entre las variables. Como resultado, se obtiene la matriz de decisión normalizada N , en la cual cada valor se ajusta a una escala común sin alterar las relaciones fundamentales entre los datos, véase ecuación 4.11.

$$N = \begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1r} \\ n_{21} & n_{22} & \cdots & n_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ n_{n1} & n_{n2} & \cdots & n_{nr} \end{pmatrix} \quad (4.11)$$

Donde cada elemento n_{ij} representa el valor normalizado, Para $i = 1, \dots, n$ y $j = 1, \dots, r$. En este los renglones están representados por las variables.

Esta matriz sirve como insumo para el cálculo de la entropía y la ponderación de variables.

Paso 5. Cálculo de la Entropía

En este paso se calcula la entropía de Shannon a la matriz de datos normalizada N con la ecuación 4.12 (Saraswat & Digalwar, 2021).

$$e_{ij} = -\frac{1}{\ln(n)} \sum_{k=1}^n \frac{n_{ik}}{n} \ln\left(\frac{n_{ik}}{n}\right) \quad (4.12)$$

Donde:

e_{ij} es la entropía asociada al renglón i y la columna j .

n_{ik} es el valor normalizado del renglón i en la columna j .

n es el número total de renglones.

$\frac{n_{ik}}{n}$ es la proporción del valor n_{ik} en relación al total de los valores en la columna.

Posteriormente, se calcula la entropía de la columna j mediante la ecuación 4.13.

$$H_j = \frac{1}{n} \sum_{i=1}^n e_{ij} \quad (4.13)$$

Donde:

H_j es la entropía de la columna j .

n es el número total de renglones de la matriz N .

e_{ij} es la entropía asociada al renglón i y la columna j .

Paso 6. Obtención de los pesos de las variables

Los pesos de las variables se determinan a partir de la entropía calculada, como se muestra en la ecuación 4.14.

$$W_j = \frac{1-H_j}{n-\sum_{k=1}^n H_k} \quad (4.14)$$

Donde:

W_j es el valor del peso de la variable j

H_j es la entropía asociada a la variable j

n es el número total de variables.

En este proceso, se calcula un peso para cada variable, el cual se interpreta como la importancia relativa de la variable en el modelo. Los pesos están normalizados en un rango de 0 a 1, lo que permite identificar el nivel de importancia de cada variable dentro del conjunto de datos.

El resultado genera los pesos calculados son presentados en un ranking de valores, que refleja la influencia de cada variable en el análisis global. Las variables con pesos

más altos indican una mayor influencia sobre el modelo, mientras que las variables con pesos bajos tienen una menor contribución.

4.2.3. Caso de aplicación del método de ponderación de variables PCA-Entropía

La implementación y validación del método de ponderación de variables propuesto, PCA-Entropía, se llevó a cabo dentro del marco de la Toma de Decisiones Multicriterio, utilizando el caso de la marginación en México por localidad, a través de su índice. El índice de marginación por localidad es una medida que evalúa el grado de rezago social y económico de una zona, considerando variables como acceso a servicios básicos, educación, vivienda e ingresos. Este índice permite clasificar las localidades en diferentes niveles de marginación, ayudando a identificar áreas con mayores necesidades de intervención (CONAPO, 2022).

La validación se aplicó específicamente al caso de estudio, con el fin de garantizar que los pesos asignados a cada variable fueran adecuados y reflejaran correctamente las disparidades de los niveles de marginación entre las localidades; con respecto de los datos reportados en las fuentes oficiales. La metodología implementada se basó en la metodología de la Toma de decisiones Multicriterio, descrita en la sección 3.1.2.

Etapas 1. Recolección y procesamiento de los datos

Para esta investigación, se utilizó la información de la Ciudad de México, específicamente de las localidades, que incluyen colonias y pueblos. Los datos correspondientes al índice de marginación fueron extraídos de (CONAPO, 2022); mientras que los datos actualizados de las localidades se obtuvieron de la Agencia Digital de Innovación Pública del Gobierno de la Ciudad de México (ADIP CDMX, 2024).

Para el preprocesamiento de los datos, se realizó una intersección entre los conjuntos de datos extraídos para obtener la información más actualizada. También, se realizó la corrección de datos erróneos, la homogenización de los valores y se completaron manualmente algunos datos faltantes empleando la técnica de interpolación promedio. Como resultado, se generó un conjunto de datos con 1,814 registros correspondientes a las localidades de la Ciudad de México. Este conjunto de datos incluye 8 variables, entre las cuales se considera el índice de marginación por localidad.

La Tabla 4.1 presenta la descripción detallada de cada variable que considera el índice de marginación; mientras que la Tabla 4.2 muestra un resumen estadístico de los datos de cada variable; donde se describe el número total de registros, la ausencia de valores nulos, datos inconsistentes, los valores máximos y mínimos, el rango, el valor medio, la mediana, la varianza, la desviación estándar y el coeficiente de variación.

Etapas 2. Asignación de pesos de las variables

Para la conformación de la matriz de decisión, las alternativas corresponden a las distintas localidades de la Ciudad de México; mientras que las variables son las consideradas en el análisis, que incluyen el índice de marginación relacionadas con las características socioeconómicas y demográficas. Cada localidad se evalúa en función de los valores correspondientes a cada una de las variables; conformando la matriz de

decisión que será utilizada en el proceso de análisis multicriterio. La Tabla 4.3 un subconjunto de datos de la matriz de decisión.

Asimismo, las variables en este conjunto de datos se encuentran en diferentes escalas, esto puede dificultar la comparación de las variables. En este caso se aplicó la técnica de normalización Z-score, que ajusta los valores de las a una media cero y una desviación estándar uno. La fórmula matemática utilizada para calcular el Z-score se presenta en la ecuación 4.5. La transformación de los datos se realizó utilizando el lenguaje de programación R, mediante la función scale del paquete base. La Tabla 4.4 muestra un subconjunto de los datos con las variables transformadas.

$$z = \frac{x - \mu}{\sigma} \quad (4.15)$$

Donde

z es el valor transformado.

x es el valor original de los datos.

μ es la media de los datos de la variable.

σ es la desviación estándar de los datos de la variable.

Tabla 4.1. Asignación de pesos de las variables del índice de marginación

Variable	Descripción	Tipo de dato
ID	Id de la localidad	-
x_1	Porcentaje de la población de 15 años o más analfabeta	Porcentaje
x_2	Porcentaje de la población de 15 años o más sin educación básica	Porcentaje
x_3	Porcentaje de viviendas sin acceso a drenaje ni excusado	Porcentaje
x_4	Porcentaje de viviendas que no cuentan con energía eléctrica	Porcentaje
x_5	Porcentaje de viviendas sin acceso a agua entubada	Porcentaje
x_6	Porcentaje de viviendas que tienen piso de tierra	Porcentaje
x_7	Porcentaje de viviendas con hacinamiento	Porcentaje
x_8	Porcentaje de viviendas que no cuentan con refrigerador	Porcentaje

Tabla 4.2. Resumen estadístico del conjunto de datos

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Total de registros	1814	1814	1814	1814	1814	1814	1814	1814
Valores nulos	0	0	512	407	221	32	0	2
Datos inconsistentes	0	0	0	0	0	0	0	0
Mínimo	0.030	2.163	0	0	0	0	0.411	0
Máximo	25.142	44.921	2.956	0.868	89.320	9.490	47.295	37.126
Rango	25.111	42.758	2.956	0.868	89.320	9.490	46.884	37.126
Suma	9726.742	30816.523	86.049	59.052	1657.829	975.604	36226.684	8502.615
Mediana	4.963	17.479	0.004	0.010	0.049	0.310	21.379	3.862
Media	5.362	16.988	0.047	0.033	0.914	0.538	19.971	4.687
SE.mean	0.059	0.170	0.003	0.001	0.116	0.018	0.236	0.093
CI.mean.0.95	0.116	0.333	0.007	0.003	0.227	0.035	0.463	0.182
Varianza	6.400	52.357	0.020	0.004	24.341	0.573	100.972	15.678
Desviación estándar	2.530	7.236	0.142	0.062	4.934	0.757	10.048	3.960
Coefficiente de variación	0.472	0.426	3.003	1.908	5.398	1.408	0.503	0.845

Tabla 4.3. Subconjunto de datos de la matriz de decisión

ID	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	2.931	16.378	0.001	0.000	0.079	0.098	24.966	3.506
2	2.964	16.537	0.000	0.000	0.042	0.048	29.071	4.366
3	3.587	10.531	0.048	0.018	0.103	0.271	14.337	1.545
4	3.067	14.658	0.010	0.007	0.496	0.060	22.528	2.612
5	7.108	12.231	0.010	0.140	0.381	0.365	18.308	2.738
6	5.502	8.199	0.024	0.004	0.052	0.235	8.925	0.945
7	7.628	23.040	0.501	0.039	1.255	0.196	28.622	3.839
8	5.077	17.175	0.000	0.058	0.030	0.003	23.671	3.612
9	2.126	9.916	0.000	0.003	0.394	0.150	14.814	1.732
10	4.793	17.038	0.000	0.001	0.247	0.324	21.795	3.266
11	4.134	9.604	0.000	0.000	0.000	0.000	6.867	0.854
12	4.134	9.604	0.000	0.000	0.000	0.000	6.867	0.854
13	2.756	12.736	0.003	0.003	0.378	0.110	19.608	2.433
14	3.454	10.979	0.008	0.008	0.677	0.101	17.663	1.906
15	2.466	17.091	0.093	0.070	0.016	0.317	21.586	5.107
16	4.450	9.010	0.000	0.002	0.000	0.238	11.565	1.391
17	4.604	11.675	0.008	0.041	0.312	0.247	16.403	1.976
18	11.795	12.022	0.000	0.000	0.000	0.106	14.508	1.106
19	4.709	17.218	0.004	0.002	0.131	0.031	18.384	1.944
20	8.269	20.043	0.000	0.000	0.018	0.019	11.797	1.116

Tabla 4.4. Conjunto de datos transformado

ID	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	-0.961	-0.084	-0.324	-0.524	-0.169	-0.581	0.497	-0.298
2	-0.948	-0.062	-0.333	-0.524	-0.177	-0.647	0.906	-0.081
3	-0.702	-0.892	0.007	-0.238	-0.164	-0.352	-0.561	-0.794
4	-0.907	-0.322	-0.264	-0.412	-0.085	-0.631	0.254	-0.524
5	0.690	-0.657	-0.263	1.737	-0.108	-0.228	-0.165	-0.492
6	0.056	-1.215	-0.162	-0.458	-0.175	-0.399	-1.099	-0.945
7	0.896	0.836	3.187	0.098	0.069	-0.452	0.861	-0.214
8	-0.113	0.026	-0.333	0.415	-0.179	-0.707	0.368	-0.272
9	-1.279	-0.977	-0.333	-0.483	-0.105	-0.512	-0.513	-0.746
10	-0.225	0.007	-0.332	-0.516	-0.135	-0.282	0.182	-0.359
11	-0.486	-1.020	-0.333	-0.523	-0.185	-0.710	-1.304	-0.968
12	-0.486	-1.020	-0.333	-0.523	-0.185	-0.710	-1.304	-0.968
13	-1.030	-0.588	-0.313	-0.477	-0.109	-0.565	-0.036	-0.569
14	-0.754	-0.830	-0.274	-0.389	-0.048	-0.576	-0.230	-0.702
15	-1.145	0.014	0.318	0.607	-0.182	-0.292	0.161	0.106
16	-0.361	-1.103	-0.333	-0.497	-0.185	-0.396	-0.836	-0.833
17	-0.300	-0.734	-0.278	0.143	-0.122	-0.385	-0.355	-0.685
18	2.543	-0.686	-0.333	-0.524	-0.185	-0.570	-0.544	-0.904
19	-0.258	0.032	-0.306	-0.493	-0.159	-0.669	-0.158	-0.693
20	1.149	0.422	-0.333	-0.524	-0.182	-0.686	-0.813	-0.902

En esta etapa se utilizó el método PCA-Entropía propuesto para determinar los valores de los pesos al conjunto de datos de marginación. A continuación, se describen los pasos del método de ponderación propuesto aplicado al caso de estudio de la marginación.

Paso 1. Cálculo de la matriz de covarianza.

En este paso se calculó la matriz de covarianza. La matriz de covarianza captura las relaciones lineales entre las variables del conjunto de datos. La Tabla 4.5 muestra el resultado de la matriz de covarianza.

Tabla 4.5. Matriz de covarianza

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1.000	0.009	-0.047	0.073	-0.045	0.008	0.015	-0.041
x_2	0.009	1.000	0.212	0.298	0.251	0.492	0.930	0.811
x_3	-0.047	0.212	1.000	0.188	0.296	0.251	0.233	0.292
x_4	0.073	0.298	0.188	1.000	0.303	0.425	0.323	0.417
x_5	-0.045	0.251	0.296	0.303	1.000	0.427	0.249	0.431
x_6	0.008	0.492	0.251	0.425	0.427	1.000	0.494	0.600
x_7	0.015	0.930	0.233	0.323	0.249	0.494	1.000	0.828
x_8	-0.041	0.811	0.292	0.417	0.431	0.600	0.828	1.000

Paso 2. Cálculo de los valores y vectores propios

A partir de la matriz de covarianza, se calcularon los valores propios y los vectores propios. Los valores propios indican la magnitud de la varianza explicada por cada componente principal. Los vectores propios representan las direcciones de máxima varianza en los datos. Estos vectores propios permiten proyectar los datos en un espacio de menor dimensión. Los cálculos se realizaron con el lenguaje R y la librería FactoMineR. La Tabla 4.6 muestra los valores propios extraídos y el porcentaje de varianza explicada correspondiente a cada componente principal. Por otro lado, la Tabla 4.7 muestra la matriz vector propios.

Tabla 4.6. Vector de valores propios

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
0.459	0.140	0.129	0.100	0.082	0.060	0.021	0.009
0.459	0.599	0.728	0.828	0.910	0.970	0.991	1.000

Tabla 4.7. Matriz de vectores propios

CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8
0.003	-0.274	0.888	0.313	0.187	-0.002	0.049	0.006
-0.451	-0.377	-0.140	0.104	0.022	-0.119	-0.393	0.675
-0.216	0.502	-0.051	0.774	-0.309	0.062	-0.019	0.016
-0.289	0.239	0.376	-0.453	-0.656	-0.278	-0.074	0.017
-0.281	0.546	0.090	-0.140	0.651	-0.384	-0.156	-0.023
-0.386	0.165	0.120	-0.235	0.109	0.856	-0.098	-0.015
-0.457	-0.364	-0.127	0.113	-0.016	-0.129	-0.273	-0.734
-0.479	-0.123	-0.094	-0.009	0.062	-0.090	0.854	0.072

Paso 3. Obtención de la matriz de componentes principales

Para asegurar una representación adecuada de la variabilidad inherente en los datos, se estableció un umbral del 95% de varianza explicada. Esta decisión fundamenta la selección de los primeros cinco vectores de la matriz de vectores propios. El resultado es una matriz de decisión de 8×5. La Tabla 4.8 ofrece una visualización clara de esta matriz.

Tabla 4.8. Matriz de componentes principales

CP1	CP2	CP3	CP4	CP5
0.003	-0.274	0.888	0.313	0.187
-0.451	-0.377	-0.140	0.104	0.022
-0.216	0.502	-0.051	0.774	-0.309
-0.289	0.239	0.376	-0.453	-0.656
-0.281	0.546	0.090	-0.140	0.651
-0.386	0.165	0.120	-0.235	0.109
-0.457	-0.364	-0.127	0.113	-0.016
-0.479	-0.123	-0.094	-0.009	0.062

Paso 4. Normalización de la matriz de componente principales

Generalmente, la normalización de la matriz en el método de entropía consiste en dividir cada elemento por la suma total de la columna correspondiente, de modo que la suma por columna sea igual a uno. Sin embargo, debido a las propiedades matemáticas de ortonormalidad de la matriz de vectores propios; donde la norma de cada componente es uno y los vectores son independientes, en este caso el proceso consistió en elevar al cuadrado cada elemento, obteniendo así elementos normalizados dentro del rango [0,1]. La Tabla 4.9 muestra la matriz de componentes principales normalizada.

Tabla 4.9. Matriz de componentes principales normalizada

CP1	CP2	CP3	CP4	CP5
0.000009	0.075076	0.788544	0.097969	0.034969
0.203401	0.142129	0.0196	0.010816	0.000484
0.046656	0.252004	0.002601	0.599076	0.095481
0.083521	0.057121	0.141376	0.205209	0.430336
0.078961	0.298116	0.0081	0.0196	0.423801
0.148996	0.027225	0.0144	0.055225	0.011881
0.208849	0.132496	0.016129	0.012769	0.000256
0.229441	0.015129	0.008836	0.000081	0.003844

Paso 5. Cálculo de la Entropía

El siguiente paso es calcular la entropía de la matriz de componentes; la cual permite evaluar la información y la incertidumbre presentes en el conjunto de datos. La Tabla 4.10 muestra el vector de los valores de entropía para cada componente.

Tabla 4.10. Cálculo de la Entropía de cada componente

CP1	CP2	CP3	CP4	CP5
0.183	0.185	0.221	0.205	0.207

Paso 6. Obtención de los pesos de las variables

Este proceso concluye es el cálculo de los pesos de las distintas variables. La Figura 4.4 muestra los pesos obtenidos para cada una de las variables a través de un ranking. Los valores de los pesos oscilan entre 0 y 1. Esto permite observar claramente el orden de importancia de las variables dentro del conjunto de datos. Dichos pesos serán los coeficientes del modelo multicriterio para obtener el nivel de marginación de las localidades.

Los resultados muestran que el atributo de mayor importancia corresponde a x_1 , asociado con el porcentaje de la población de 15 años y más que es analfabeta. Asimismo, destaca el atributo x_3 , vinculado al porcentaje de viviendas que no disponen de drenaje o servicios sanitarios, el cual presenta también un peso significativo. Por el contrario, el atributo con menor relevancia es x_8 , relacionado con el porcentaje de viviendas que carecen de refrigerador.

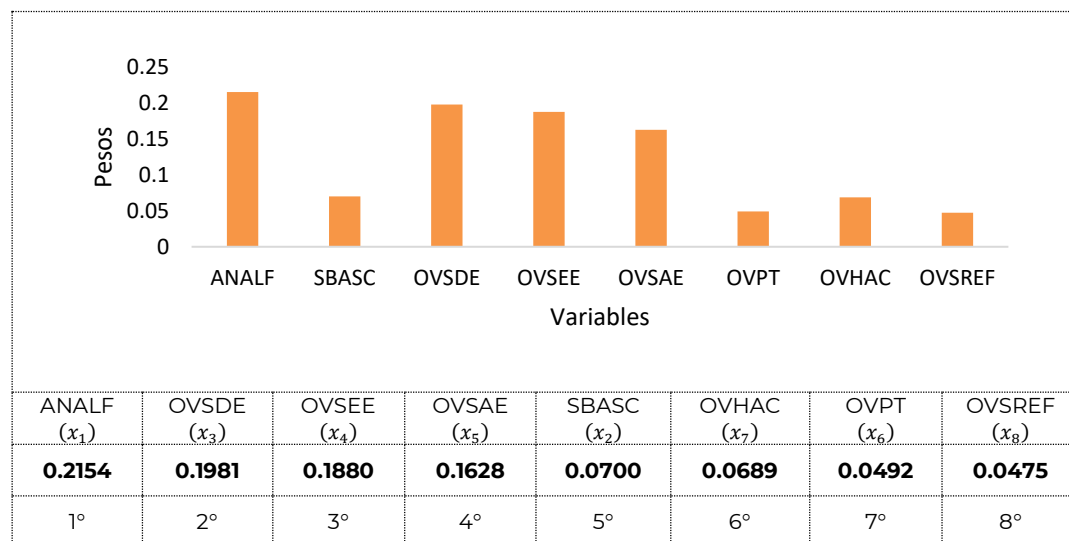


Figura 4.4. Pesos obtenidos mediante el método PCA-Entropía

Etapas 3. Evaluación de las alternativas

Una vez obtenidos los pesos, se formuló el modelo matemático multicriterio, basado en una función de agregación exponencial. Esta función fue seleccionada debido a su capacidad para capturar mejor las diferencias en los niveles de marginación; permitiendo una representación más precisa de las disparidades entre localidades. En este modelo, los coeficientes de la función de agregación corresponden a los valores

de los pesos obtenidos mediante el método PCA-Entropía. La ecuación del modelo de agregación exponencial 4.16 utilizada.

$$f(x) = e^{0.2154x_1} + e^{0.0700x_2} + e^{0.1981x_3} + e^{0.1880x_4} + e^{0.1628x_5} + e^{0.0492x_6} + e^{0.0689x_7} + e^{0.0475x_8} \quad (4.16)$$

A partir del modelo formulado, se evaluaron las distintas alternativas del conjunto de datos mediante el Modelo Multicriterio de Agregación Exponencial, obteniendo un puntaje global para cada una de ellas. El resultado final permitió establecer un ranking que ordena las alternativas de acuerdo con su desempeño, desde la menor hasta la de mayor preferencia.

Para la selección de alternativas, se utilizó el método de optimización de intervalos de Jenks, que minimiza la varianza dentro de cada grupo y maximiza la diferencia entre los grupos. Este método es particularmente eficaz para segmentar datos, garantizando que los grupos sean lo más homogéneos posible internamente y distintos entre sí. Los resultados obtenidos muestran un subconjunto que clasifica las localidades en 5 grupos con base a su nivel de marginación; distribuidas en cinco categorías: (1) Muy baja, (2) Baja, (3) Media, (4) Alta y (5) Muy alta. La Tabla 4.11 muestra algunos datos de localidades y el nivel de marginación.

Tabla 4.11. Subconjunto de datos clasificado

ID	ANALF	SBASC	OVSDE	OVSEE	OVSSE	OVPT	OVHAC	OVSREF	Datos evaluados	Datos clasificados
609	-0.074	0.507	-0.088	-0.126	-0.176	-0.451	1.227	0.620	1.438	3
610	1.112	0.142	-0.333	0.477	-0.166	-0.710	-0.171	0.795	1.145	3
611	-0.074	0.507	-0.088	-0.126	-0.176	-0.451	1.227	0.620	1.438	3
612	-0.629	-0.848	-0.333	-0.505	-0.185	-0.644	-1.245	-0.895	-5.285	1
613	3.295	0.420	-0.162	0.067	-0.174	-0.135	0.812	0.075	4.200	3
614	1.205	-0.985	-0.303	-0.504	-0.185	-0.142	-1.212	-0.819	-2.946	1
615	-0.701	1.181	-0.192	0.115	-0.160	-0.242	0.933	0.540	1.473	3
616	-0.239	0.023	-0.333	-0.524	-0.134	-0.582	0.532	-0.064	-1.322	2
617	-0.252	-0.209	0.070	0.054	-0.131	-0.120	0.343	-0.068	-0.314	2
618	-0.452	1.444	0.121	0.182	-0.182	0.929	1.872	1.568	5.481	3
619	-1.392	-0.004	2.209	-0.524	-0.159	-0.232	-0.460	-0.810	-1.372	2
620	-0.581	-0.921	-0.333	-0.523	-0.185	-0.243	-1.044	-0.757	-4.587	1
621	-0.634	-0.563	-0.317	3.478	-0.166	-0.395	-0.037	-0.252	1.115	3
622	-0.985	2.128	-0.332	-0.515	-0.185	0.327	1.911	2.146	4.496	3
623	-0.545	0.212	-0.226	-0.521	-0.168	-0.519	0.332	-0.130	-1.565	2
624	-0.192	-0.156	-0.309	-0.524	-0.181	-0.270	-0.165	-0.308	-2.106	2
625	-0.192	-0.156	-0.309	-0.524	-0.181	-0.270	-0.165	-0.308	-2.106	2
626	-0.157	0.633	-0.306	-0.182	-0.178	-0.019	0.960	0.370	1.121	3
627	-0.157	0.633	-0.306	-0.182	-0.178	-0.019	0.960	0.370	1.121	3

Etapas 4. Validación

Como variable de referencia se utilizó el Índice de Marginación, elaborado por el Consejo Nacional de Población (CONAPO). Este indicador, de carácter oficial, constituye una medida consolidada y ampliamente reconocida que refleja de manera integral las carencias sociales, con un nivel de desagregación a escala de localidad. El índice sintetiza múltiples dimensiones asociadas a carencias en educación, condiciones de la vivienda y acceso a bienes y servicios, lo que lo convierte en un referente adecuado para evaluar y contrastar los resultados generados por el modelo propuesto.

En esta etapa se emplearon dos enfoques complementarios para evaluar la precisión y validez del modelo multicriterio desarrollado. El primer enfoque consistió en aplicar la correlación de Spearman con el objetivo de medir la relación y consistencia entre los valores generados por el modelo y los valores del Índice de Marginación. Esta técnica no paramétrica fue seleccionada debido a que no asume una distribución normal en los datos y es especialmente útil cuando se desea comparar la consistencia entre ambos modelos. En este caso, permitió evaluar si existe una asociación significativa entre el modelo y el proporcionado por el índice de marginación.

El segundo enfoque consistió en utilizar una matriz de confusión, que permite evaluar el desempeño del modelo como un clasificador. Para ello, se comparó la clasificación de las localidades generada por el modelo multicriterio y los datos de referencia del índice de marginación. Esta matriz proporciona información detallada sobre el número de coincidencias y discrepancias entre ambas clasificaciones. La elección de esta herramienta se justifica porque ofrece una validación categórica directa y útil para evaluar el grado de concordancia entre las decisiones del modelo y los datos de referencia.

Los resultados muestran una alta correlación positiva, lo que indica que el modelo multicriterio logra capturar de manera adecuada los patrones de coherencia y concordancia entre las localidades. Esto evidencia un fuerte grado de asociación entre los valores generados por el modelo y los datos oficiales, como se muestra en la Tabla 4.12. Asimismo, el análisis de la matriz de confusión permitió evaluar la precisión del modelo al identificar el grado de coincidencia entre las categorías asignadas por el modelo y las categorías reales establecidas por el índice de referencia, como se presenta en la Tabla 4.13.

La combinación de la matriz de confusión con el coeficiente de Spearman, permite obtener un análisis robusto al evaluar de manera integral el desempeño del modelo de clasificación del índice de marginación con enfoque multicriterio en relación con un índice de referencia. Se observa que las clases más representadas, como la Clase 1 y la Clase 2, presentan mayor precisión; mientras que las clases con menor cantidad de datos, como la Clase 4 y la Clase 5 muestran una precisión reducida y una tendencia a la clasificación errónea en otras categorías. Además, se destaca que la Clase 2 es frecuentemente confundida con la Clase 3. Esto sugiere que las variables de marginación en estos niveles pueden ser similares.

Por otro lado, el coeficiente de correlación de Spearman, con un valor de $\rho = 0.8464$ y un p-valor de 0, confirma que existe una relación fuerte y significativa entre el índice

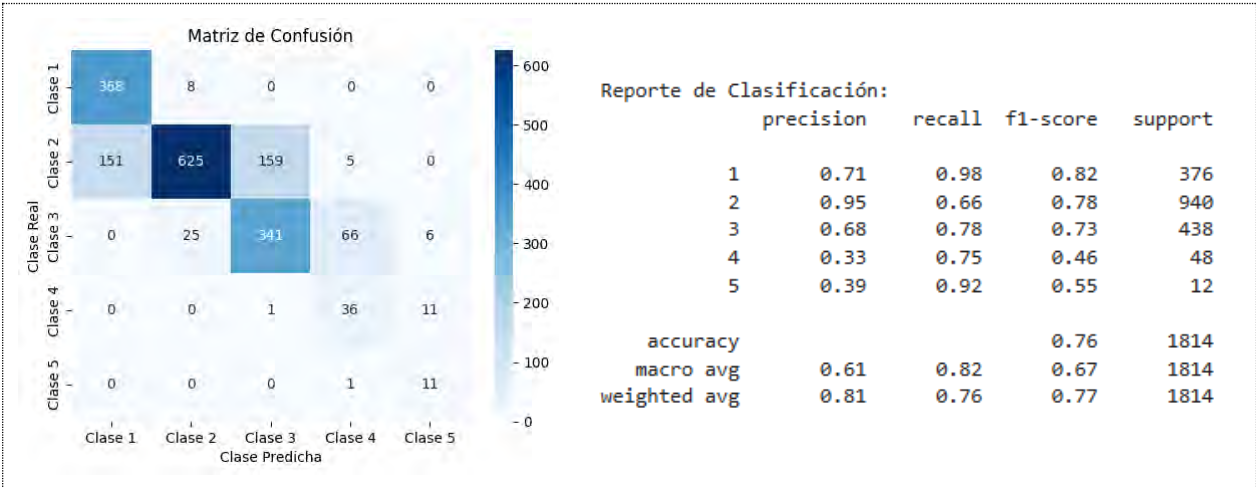
de marginación obtenido por el modelo y el índice de referencia. Este resultado respalda la coherencia global de la clasificación, indicando que, a pesar de los errores específicos detectados en la matriz de confusión, el modelo sigue una tendencia similar al índice tradicional y conserva la estructura ordinal de las categorías de marginación.

La combinación de estos enfoques proporciona una evaluación robusta del modelo. Mientras que la matriz de confusión permite identificar errores concretos y áreas de mejora, la correlación de Spearman valida que la clasificación general es consistente con la referencia. En conclusión, la matriz de confusión y el coeficiente de Spearman se complementan para proporcionar un diagnóstico detallado y una validación estadística del modelo de marginación, permitiendo ajustes estratégicos que mejoren su desempeño en futuras aplicaciones. Los resultados confirmaron que el método PCA-Entropía genera una clasificación consistente con el índice de marginación, asegurando su aplicabilidad como una herramienta alternativa en el contexto de la ponderación de variables. Este enfoque metodológico permitió una evaluación objetiva y estructurada de los niveles de marginación, integrando técnicas avanzadas de ponderación y clasificación para obtener resultados más precisos y representativos.

Tabla 4.12. Resultado del coeficiente de correlación de Spearman

Datos de referencia del índice de marginación por localidad vs Índice de marginación por localidad obtenido con enfoque multicriterio	$\rho=0.8464144719910918$ p-valor: 0
---	---

Tabla 4.13. Resultado de la matriz de confusión



4.3. Metodología de solución para el desarrollo del modelo matemático multicriterio

En esta investigación, se adoptó un marco estructurado basado en la metodología de análisis de decisiones multicriterio, comúnmente conocida como el Proceso de Toma de Decisiones Multicriterio (MCDM, por sus siglas en inglés). Esta metodología se centra en seleccionar la mejor alternativa de un conjunto de opciones, tomando en

cuenta diversas variables. Su aplicación es especialmente valiosa cuando las decisiones no pueden basarse en una sola variable, sino que requieren la evaluación de múltiples dimensiones, incluidas aquellas que son intangibles o difíciles de cuantificar (Zarghami & Szidarovszky, 2011), (Taherdoost & Madanchian, 2023), (Zardari et al., 2015) y (Grajales et al., 2013).

Asimismo, se consultaron dos metodologías complementarias: CRISP-DM para la Ciencia de Datos y la metodología para la construcción de indicadores. CRISP-DM es una de las metodologías ampliamente utilizadas en proyectos de minería de datos y ciencia de datos, debido a su enfoque estructurado y su versatilidad para abordar una variedad de problemas. Proporciona un marco detallado que orienta cada fase del proceso de análisis de datos, desde la comprensión del problema hasta la implementación del modelo. La inclusión de CRISP-DM en este estudio es crucial, ya que asegura que el proceso de modelado y análisis de datos se realice de manera sistemática y eficiente (Espinosa, 2020), (Huber, 2019), (Wiemer et al., 2019).

La metodología para la construcción de indicadores se consultó debido a su capacidad para proporcionar una herramienta clara y accesible para la interpretación del modelo, a través de un indicador. Estos indicadores son fundamentales para evaluar y cuantificar las dimensiones del fenómeno en estudio, en este caso, la accesibilidad a los servicios. Esto mejora la precisión, fiabilidad y claridad de los resultados obtenidos, optimizando la toma de decisiones (Schuschny & Soto, 2009), (CONEVAL, 2024) y (OECD/European Union/EC-JRC, 2008).

La metodología propuesta se basa principalmente en la Toma de Decisiones Multicriterio, adaptándolo a las necesidades del contexto específico. La combinación de esta metodología con CRISP-DM y la metodología para la construcción de indicadores permite proporcionar un enfoque robusto y complementario que enriquece el proceso de desarrollo del modelo multicriterio. CRISP-DM resulta esencial para estructurar adecuadamente las fases de preprocesamiento, específicamente el análisis de los datos y la selección de variables relevantes. Por su parte, la metodología para la construcción de indicadores es clave para diseñar indicadores precisos que midan la accesibilidad a los servicios, proporcionando una representación clara y cuantificable de las diversas dimensiones del problema. Esto permite una evaluación detallada de las alternativas, enriqueciendo el análisis, mejorando la calidad de la toma de decisiones y la interpretación y comprensión del problema.

La Figura 4.5 se presenta la metodología de solución propuesta para el desarrollo del modelo matemático multicriterio de accesibilidad a los servicios; el cual está definido en tres fases y cada fase considera diversos procesos. A continuación, se describen los procesos de cada una de las fases.

Fase 1. Preprocesamiento de los datos

Etapas 1.1. Recolección de los datos

Este proceso comienza con la definición de las alternativas y las variables relacionados con el problema multicriterio, conocidos comúnmente como instancias y variables, respectivamente. A continuación, se lleva a cabo la búsqueda y recopilación de información proveniente de diversas fuentes confiables y actualizadas; asegurando

que los datos recolectados sean relevantes y adecuados para el análisis en el contexto específico del problema.

Etapa 1.2. Preparación de los datos

Este proceso incluye la limpieza de datos, que abarca corrección de errores, imputación de valores faltantes, la homogeneización de la información y la eliminación de redundancias para garantizar la calidad del conjunto de datos. Asimismo, la generación de nuevas variables a partir de los existentes, con el objetivo de enriquecer y ampliar la información disponible. También se realiza una descripción y exploración detallada de los datos para identificar patrones significativos que puedan influir en el análisis. Por último, las variables se transforman a una escala común, lo que facilita su comparación y asegura una contribución equilibrada al análisis y al modelado.

Fase 2. Construcción del modelo multicriterio

Etapa 2.1. Identificación de variables relevantes

Este proceso consiste en seleccionar las variables más significativas y pertinentes del conjunto de datos para el análisis del problema en estudio. La relevancia de cada variable se evalúa acorde a su impacto potencial en los resultados y su capacidad para aportar una comprensión más profunda del problema. Identificar de manera precisa las variables relevantes es fundamental para garantizar que el análisis se enfoque en los factores clave, mejorando la precisión y la efectividad del modelo.

Etapa 2.2. Asignación de pesos de las variables

Este proceso es fundamental dentro de la metodología; conocido como ponderación de variables, este proceso se enfoca en determinar los pesos de los coeficientes de las variables que conforman el modelo matemático multicriterio. Estos coeficientes son clave para expresar el nivel de importancia que tiene de cada variable en relación con los demás. Una ponderación adecuada es crucial para asegurar que el modelo final sea preciso, relevante y alineado con el contexto del problema en estudio.

Etapa 2.3. Formulación del modelo

El primer paso consiste en seleccionar una función matemática de agregación que permita integrar las variables relevantes identificadas en la etapa 2.1. Esta función da lugar a un modelo matemático multicriterio representado por una ecuación, que combina de manera adecuada las variables y sus respectivos pesos para reflejar la influencia de cada uno en el resultado final.

En el siguiente paso las alternativas se evalúan al sustituir sus valores correspondientes en el modelo matemático. Esto permite calcular una puntuación para cada alternativa, basada en las variables establecidos del modelo. Las puntuaciones se ordenan de menor a mayor, facilitando su comparación. Este análisis es clave para identificar la alternativa que mejor se alinea con los objetivos establecidos y tomar la decisión para seleccionar la opción más adecuada.

Fase 3: Validación e interpretación del modelo

Etapas 3.1. Análisis comparativo

Análisis comparativo consiste en realizar una comparación del modelo matemático multicriterio con otros modelos existentes para validar su fiabilidad y coherencia. Al contrastar los resultados del modelo con otros enfoques, se puede verificar la consistencia del modelo; es decir, si los resultados obtenidos son coherentes con los de otros modelos aplicados en contextos similares.

Este proceso permite identificar las fortalezas y debilidades del modelo, así como detectar patrones o tendencias comunes que puedan respaldar su validez. Asimismo, al utilizar métricas en la comparación, se puede garantizar que el modelo propuesto sea preciso, relevante y aplicable para abordar el problema en estudio. Esta comparación no solo valida el modelo en términos de sus resultados, sino que también proporciona una base sólida para posibles mejoras y ajustes.

Etapas 3.2. Análisis de sensibilidad

El análisis de sensibilidad es un proceso esencial en la Toma de Decisiones Multicriterio; ya que permite examinar la robustez y estabilidad del modelo frente a variaciones en los parámetros, como cambios en los pesos de las variables, cambios en los valores de las alternativas, entre otros. Es decir, permite identificar cómo pequeños ajustes en las variables de entrada afectan los resultados del modelo y si estos cambios alteran significativamente la solución final.

Este análisis permite asegurar que el modelo sea confiable, incluso en contextos de incertidumbre o cuando se presenten escenarios inesperados. Garantiza que el modelo conserve su efectividad y adaptabilidad en diversas condiciones; lo que aumenta la confianza en su capacidad para ofrecer soluciones válidas y óptimas en situaciones prácticas y reales.

Etapas 3.3. Interpretación del modelo

Este proceso consiste en aplicar variables específicas derivados del modelo multicriterio para generar información resumida y sintetizada, a través de la medición de su desempeño. El resultado es una herramienta que representa con precisión el fenómeno en estudio. El propósito de este enfoque es garantizar que el modelo refleje de manera adecuada los aspectos más relevantes, facilitando así una toma de decisiones más informada y efectiva.

Para un análisis más claro del desempeño, los resultados deben representarse de forma cartográfica. La visualización en mapas facilita la comprensión al mostrar la distribución geográfica de los resultados, permitiendo identificar patrones y tendencias espaciales. Esta representación visual no solo mejora la interpretación de los datos, sino que también optimiza el proceso de toma de decisiones al proporcionar una visión más clara y accesible de los resultados.

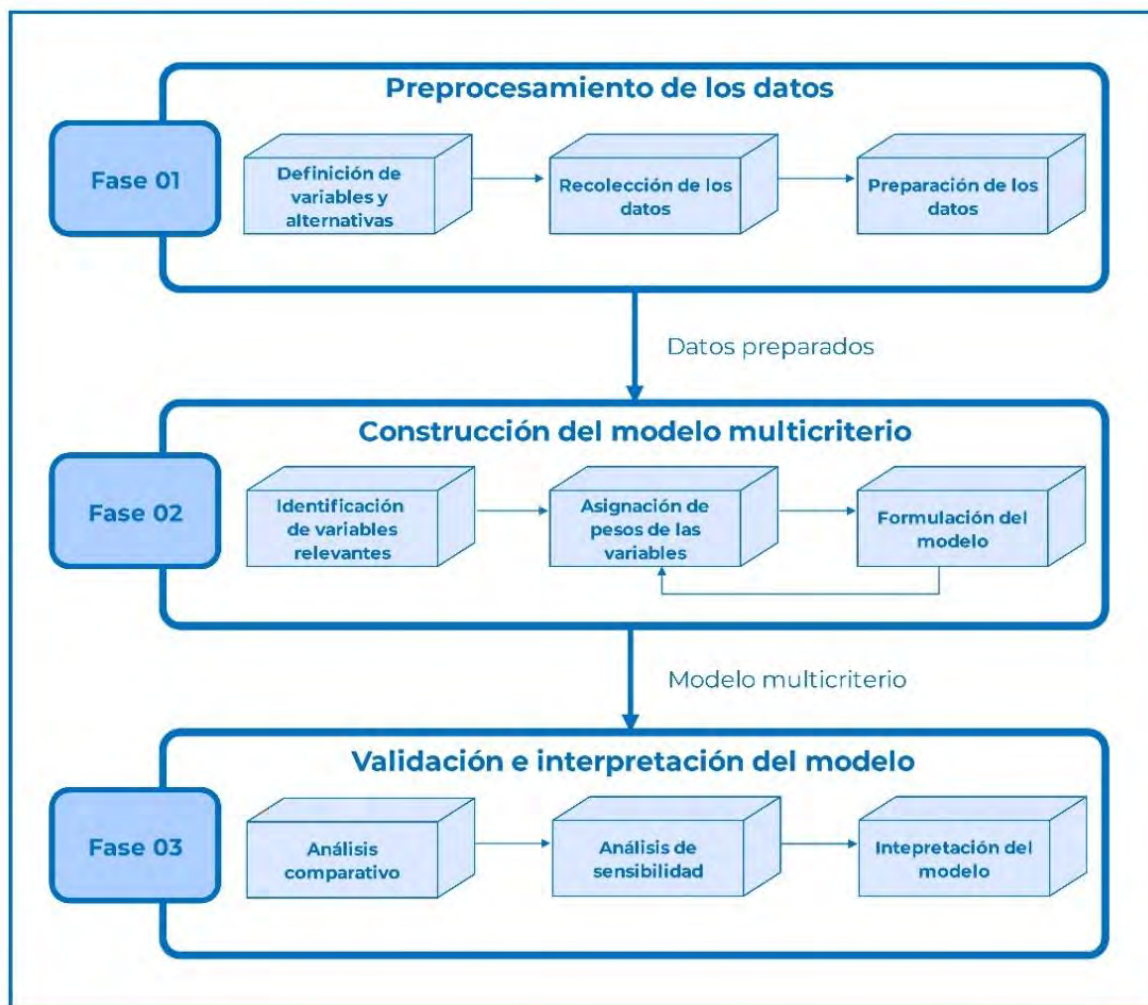


Figura 4.5. Esquema de la metodología de solución

4.3.1. Caso de estudio utilizado

Esta investigación se aplicará a un caso de estudio real relacionado con la **Accesibilidad a los Servicios**. Según lo establecido en la sección 3.6.3, la accesibilidad se entiende como la relación entre un punto de origen y un punto de destino dentro del espacio geográfico, determinada por la interacción de diversas variables sociales, demográficas, económicas y espaciales.

El análisis de la accesibilidad constituye un problema complejo que implica múltiples variables interrelacionadas, las cuales deben ser consideradas de manera simultánea. Este tipo de problema puede clasificarse como un problema bajo conflicto, ya que la mejora en una variable puede generar efectos negativos en otras. Por ejemplo, una alta densidad poblacional puede coincidir con una baja disponibilidad de escuelas, lo que refleja un desequilibrio en el acceso a los servicios. Además, estas variables no son independientes entre sí; su mutua influencia incrementa la complejidad del proceso de toma de decisiones al momento de seleccionar alternativas viables.

En consecuencia, esta investigación se orienta a identificar zonas de interés que presentan mayores dificultades de acceso a los servicios, así como aquellas con menor

nivel de afectación. Para ello, se integra el peso relativo de cada variable junto con una función matemática de agregación; con el fin de encontrar soluciones compromiso que permitan equilibrar adecuadamente los distintos factores involucrados. El objetivo es lograr una evaluación más precisa de la accesibilidad y proporcionar una base sólida para la toma de decisiones informadas en la selección de alternativas, con especial énfasis en la identificación de zonas prioritarias.

Línea base en el estudio de la accesibilidad

En el estudio de la accesibilidad, sobresalen los modelos de accesibilidad espacial enfocados en la dimensión geográfica; los cuales permiten medir el nivel de accesibilidad de forma cuantitativa. Estos modelos consideran principalmente variables como la distancia, el tiempo de viaje, el número de servicios disponibles y, en algunos casos, la población. No obstante, tienden a dejar de lado otras dimensiones relevantes, como las condiciones sociales y económicas de la población. Entre los enfoques más representativos se encuentran las medidas basadas en la distancia, el indicador de oportunidades acumuladas, el indicador de interacción espacial, el indicador de separación y el método de captación flotante (Hansen, 1959), (Ingram, 1971), (Wachs & Kumagai, 1973), (Knox, 1978), (Garrocho, C. F. & Campos, 2006).

Por otro lado, el índice de marginación por localidad constituye una medida del nivel de carencias sociales de una zona geográfica, al integrar variables como el acceso a servicios básicos, la educación, la calidad de la vivienda y los niveles de ingreso (CONAPO, 2022). Si bien este índice no evalúa directamente la accesibilidad en términos espaciales; es decir, no contempla la relación entre la ubicación de la población y la localización de los servicios, sí aporta información valiosa sobre las condiciones estructurales que influyen en el comportamiento territorial y en las necesidades sociales de las localidades. En este sentido, ofrece una perspectiva complementaria que, al combinarse con los enfoques espaciales, permite construir una visión más amplia, multidimensional y operativa del fenómeno de la accesibilidad.

Por ello, esta investigación propone una aproximación integrada y multidimensional que complementa las líneas base existentes, con el fin de avanzar hacia una comprensión más completa del fenómeno de accesibilidad y su impacto territorial.

Zona de estudio

En esta investigación, la Ciudad de México fue seleccionada como la zona geográfica de estudio. La Ciudad de México, también conocida como la capital de los Estados Unidos Mexicanos, está ubicada en el centro del país. Su extensión territorial es de 1,485 kilómetros cuadrados, lo que representa apenas el 0.1 % del territorio nacional. Desde 2016, está conformada por 16 municipios territoriales, conocidas como alcaldías. Cada municipio está constituido por diversas colonias y pueblos, definidos como localidades (INEGI, 2024).

Según el censo de 2020, la población total de la Ciudad de México asciende a 9,209,944 habitantes, lo que equivale al 7.3 % de la población nacional. De esta cifra, el 99 % corresponde a población urbana y solo el 1 % a población rural, en contraste con el promedio nacional, donde el 79 % de la población es urbana y el 21 % rural. Dentro de

la Ciudad de México, el grupo de edad de 25 a 29 años es el más representativo, con un total de 752,289 personas, lo que equivale al 8.2 % de la población total.

Otros datos relevantes de la Ciudad de México incluyen que la población económicamente activa (PEA) constituye el 70.8 % de los habitantes, siendo una de las proporciones más altas a nivel nacional. La actividad económica de la ciudad se concentra principalmente en los municipios de Cuauhtémoc, Miguel Hidalgo, Álvaro Obregón y Benito Juárez, que en conjunto representan el 23.4 % de la población total de la ciudad. Estas características destacan la relevancia económica y demográfica de la Ciudad de México dentro del contexto nacional (SEDECO CDMX, 2023), (GCDMX, 2025).

Accesibilidad a servicios básicos: escuelas de nivel primaria

Un servicio se define como un conjunto de acciones realizadas con el fin de satisfacer las necesidades de los habitantes de un lugar. Los servicios tienen el propósito de que las personas reciban lo necesario para su bienestar. Los servicios básicos en una localidad, comprenden las infraestructuras esenciales para garantizar una vida saludable. Por tal motivo, mejorar y ampliar la provisión de estos servicios debe ser un componente clave en el desarrollo de la comunidad (Padilla et al., 2019), (INEGI, 2024), (GCDMX, 2025).

En esta investigación se consideran como servicios las escuelas primarias, tanto públicas como privadas, ubicadas en diversas localidades de la Ciudad de México. La selección de estas escuelas se fundamenta en su relevancia según el artículo 3° de la Constitución Política de los Estados Unidos Mexicanos, que establece el derecho de toda persona a la educación, enfatizando la importancia de garantizar el acceso a la educación básica. Asimismo, destaca la necesidad de implementar acciones específicas para mejorar las condiciones de vida en las escuelas situadas en zonas de alta marginación.

Las escuelas primarias, pertenecientes al nivel básico, son reconocidas como servicios esenciales, pues forman parte de las Unidades Básicas de Servicio en cada región, posicionándose como elementos clave para asegurar la cobertura educativa de la población (SEDESOL, 1992). Garantizar que cada localidad cuente con este servicio resulta fundamental para fomentar la equidad y el desarrollo social. Por ello, la inclusión de las escuelas primarias en este estudio es indispensable para analizar y evaluar la accesibilidad educativa en la ciudad, ofreciendo un panorama integral sobre su distribución y las oportunidades reales de acceso para los habitantes.

En el presente estudio se priorizó el transporte peatonal como modalidad principal de acceso, dado que las escuelas de nivel básico suelen estar estratégicamente distribuidas dentro de cada localidad para facilitar el acceso sin necesidad de realizar desplazamientos prolongados.

4.4. Desarrollo del modelo matemático multicriterio

El desarrollo del modelo matemático multicriterio, enfocado en la accesibilidad a servicios, tiene como objetivo proporcionar una herramienta analítica para evaluar la

accesibilidad; permitiendo la toma de decisiones fundamentada. Este proceso implica la integración de múltiples variables relevantes, considerando factores sociales, demográficos, económicos y espaciales que influyen, siguiendo la metodología de solución. El desarrollo del modelo se concentra en las dos primeras fases de la metodología de solución: Procesamiento de los datos y Construcción del modelo multicriterio. Este modelo facilita el análisis de accesibilidad a los servicios al ofrecer un marco estructurado para la toma de decisiones, contribuyendo a la optimización de la asignación de recursos y mejorando la planificación territorial de los interesados.

4.4.1. Preprocesamiento de los datos

En esta etapa se llevó a cabo una revisión bibliográfica exhaustiva para identificar las variables que afectan la accesibilidad en una zona geográfica, así como las alternativas relacionadas, los servicios a estudiar y las modalidades de transporte a considerar. El objetivo es la construcción de un conjunto de datos enfocado en la accesibilidad a servicios para su análisis posterior.

1) Recolección de los datos

La recolección de datos consiste en la búsqueda de información relacionada con algún fenómeno de estudio en fuentes de datos; así como la integración de dicha información en un conjunto que se ajuste al contexto. Lo anterior con el propósito de que los datos obtenidos sean precisos y adecuados para el análisis. En esta investigación, la etapa de recolección de datos no solo incluyó la recopilación de información sobre la accesibilidad a servicios, sino también la identificación de los criterios para el desarrollo de un modelo matemático con enfoque multicriterio. A continuación, se describen las actividades realizadas para este proceso.

Identificación de variables para la accesibilidad a servicios:

Con base en la revisión de la literatura y en los componentes de accesibilidad mencionados, se han definido diversos variables factores sociales, demográficos, económicos y espaciales; los cuales están relacionados que impactan en la accesibilidad. Estas variables permiten analizar cómo las distintas dimensiones influyen en el acceso a la educación básica (escuelas primarias), ofreciendo una perspectiva detallada e integral sobre las condiciones actuales de la población en la Ciudad de México.

Este enfoque integral no solo facilita un análisis exhaustivo de la accesibilidad hacia las escuelas primarias; sino que también proporciona una base sólida para tomar decisiones.

- o Variables sociales: Estos incluyen aspectos como el nivel de educación, población con educación básica y población analfabeta; los cuales relacionados con el bienestar social. Dichas variables son cruciales para identificar las zonas más desfavorecidas.
- o Variables demográficas: Analizan la distribución densidad de la población en cada localidad, la población potencial que considera la proporción de grupos de edad

que más requieren el acceso a la educación básica. Estas variables ayudan a comprender las necesidades específicas de cada zona en función de su población.

- o Variables económicas: Se refieren a la capacidad de los hogares para cubrir necesidades básicas, como acceso a bienes y tecnologías. Por ejemplo, refrigerador, automóvil, televisión, celular, internet, teléfono, entre otros. Dichas variables son importantes para evaluar las brechas económicas que impactan la accesibilidad.
- o Variables espaciales: Incluyen la ubicación de las escuelas, cantidad de escuelas en una zona geográfica, las distancias recorridas para llegar a ellas y las características de la infraestructura vial de alguna modalidad de transporte. Estas variables permiten evaluar cómo la distribución geográfica de las escuelas afecta el acceso de la población.

Identificación de alternativas para la accesibilidad a servicios

En esta investigación, se tomaron en cuenta las localidades para analizar la accesibilidad a las escuelas primarias. Como punto de referencia, se consideró el centro geográfico de cada localidad, también conocido como centroide; el cual actúa como el punto de partida común para la población que habita en el lugar. El uso de las localidades y sus centroides como puntos de referencia ofrece una base sólida para analizar y comparar la accesibilidad a las escuelas primarias en la Ciudad de México. Este enfoque no solo simplifica el análisis técnico, sino que también permite obtener resultados relevantes y representativos. La elección de las localidades como unidades de análisis se fundamenta en varias razones:

- o Representatividad geográfica: Las localidades, al ser divisiones administrativas bien delimitadas, permiten capturar características específicas de cada área, como la densidad poblacional, las condiciones socioeconómicas y la distribución de infraestructura. Esto facilita un análisis detallado y contextualizado de la accesibilidad en función de las condiciones locales. Las localidades son unidades reconocidas en la planificación y gestión de servicios públicos, lo que facilita la interpretación de los resultados y su aplicación.
- o Disponibilidad de datos: Los datos demográficos, sociales y económicos suelen estar organizados y disponibles a nivel de localidades, lo que permite integrar la información de diferentes fuentes.
- o Proximidad poblacional: El centroide de cada localidad representa un punto de referencia que, aunque no refleja la ubicación exacta de cada habitante, proporciona una aproximación razonable para calcular distancias y tiempos de acceso. Este enfoque es particularmente útil en localidades donde la población tiende a concentrarse en torno al centro.
- o Simplificación del análisis: Utilizar el centroide como punto de partida común simplifica el análisis geoespacial, ya que reduce la complejidad de trabajar con múltiples puntos de origen. Esto permite realizar comparaciones entre localidades de manera más eficiente.

Con base a lo anterior, el análisis de accesibilidad a las escuelas primarias, en esta investigación, considera la ubicación geográfica de los centros educativos como puntos de destino, mientras que los puntos de origen se establecen en el centroide de cada localidad, representando el punto común de partida para la población. Las variables que se toman en cuenta es el uso predominante del transporte peatonal como el medio principal para llegar a estos destinos. También, se incluyen variables que se relacionan con factores sociales, demográficos, económicos y espaciales que aportan una comprensión más integral de las condiciones de acceso educativo en la Ciudad de México. La Figura 4.6 muestra un resumen de los elementos para el desarrollo del modelo multicriterio, enfocado en la accesibilidad a servicios.



Figura 4.6. Elementos para el caso de estudio

Obtención de datos

Para obtener información relevante, se llevó a cabo una búsqueda en diversas fuentes de datos, priorizando bases de datos especializadas provenientes de fuentes oficiales. Posteriormente, se procedió a una minuciosa recopilación de datos que abarcó tanto información espacial, como datos no espaciales relacionados con el tema de estudio. La Tabla 4.14 proporciona una descripción detallada de las fuentes consultadas y los datos extraídos.

Tabla 4.14. Descripción de las fuentes consultadas y la información extraída

Datos recolectados	Fuente de datos	Link de acceso	Formato	Número de registros obtenidos
Municipio	Portal de Datos Abiertos del Gobierno de la Ciudad de México	https://datos.cdmx.gob.mx/dataset/alcaldias	Información espacial: Shape	16
Localidades	Portal de Datos Abiertos del Gobierno de la Ciudad de México	https://datos.cdmx.gob.mx/dataset/catalogo-de-colonias-datos-abiertos	Información espacial: Shape	1,814
Escuelas de nivel primaria: públicas y privadas	Portal de Datos Abiertos del Gobierno de la Ciudad de México	https://datos.cdmx.gob.mx/dataset/escuelas-privadas https://datos.cdmx.gob.mx/dataset/escuelas-publicas	Información espacial: csv	2,234
	Directorio Estadístico Nacional de Unidades Económicas (DENUE), Instituto Nacional de Estadística y Geografía (INEGI).	https://www.inegi.org.mx/app/mapa/denue/default.aspx	Información espacial: csv	2,382
Variables sociales, demográficas y económicas	Consejo Nacional de Población (CONAPO)	https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372	Información espacial y no espacial: Shape y csv	2,243
	Centro de Investigación en Ciencias de Información	http://pgot.centrogeo.org.mx/layers/geonode%3Amanzanas_censo2020	Información espacial y no espacial: Shape y csv	69,789 (manzanas)

También se llevó a cabo un proceso exhaustivo de limpieza de datos, que incluyó la eliminación de datos erróneos, la homogenización de unidades y valores, y la depuración de atributos con información redundante. Estas acciones fueron

realizadas de forma manual para garantizar la coherencia y calidad de los datos utilizados.

Integración del conjunto de datos

La integración del conjunto de datos se realizó mediante un proceso de intersección espacial y no espacial, utilizando herramientas como Microsoft Excel, el software de Sistemas de Información Geográfica (SIG) QGIS versión 3.34.1 y el lenguaje de programación Python. Esta etapa consistió en combinar de manera eficiente múltiples fuentes de información, asegurando su correcta alineación y comparabilidad para el análisis posterior. Este enfoque garantizó la precisión y adecuación de los datos empleados en la evaluación de la accesibilidad.

El procedimiento implicó la superposición de capas de información relevantes para generar un conjunto de datos unificado y coherente. Entre las capas consideradas se incluyeron:

- o Localidades: Representadas mediante polígonos que delimitan las zonas de origen y análisis.
- o Escuelas primarias: Incluyendo tanto instituciones públicas como privadas, consideradas como destinos educativos clave.

Variables demográficas, sociales y económicas: Criterios que permiten contextualizar y enriquecer el análisis.

Como parte del proceso, se generaron los centros geográficos de las localidades, conocidos como centroides, los cuales se utilizaron como puntos de origen en el análisis espacial. Esta estrategia permitió representar con mayor precisión y consistencia las características espaciales de interés. Los centroides fueron calculados en QGIS mediante la herramienta "Centroides", disponible en el menú Vectorial, dentro de las Herramientas de Geometría del marco de procesamiento.

La Figura 4.7 muestra un mapa con la distribución geográfica de municipios, localidades y manzanas en la Ciudad de México, donde las manzanas permiten observar la localización de las viviendas y la concentración poblacional. La Figura 4.8 presenta la ubicación geográfica de las escuelas de nivel primaria. Esta información se complementa con la Tabla 4.5 que detalla la cantidad de escuelas por municipio, proporcionando una visión más clara de su distribución espacial.

Esta etapa aseguró que el conjunto de datos final estuviera optimizado y centrado en las variables relevantes para el análisis. El conjunto resultante está compuesto por 1,811 registros o instancias, cada una asociada a una localidad dentro de la Ciudad de México. Cada instancia cuenta con 14 variables numéricas, correspondientes a dimensiones sociales, demográficas, económicas y espaciales.

Tabla 4.15. Número de escuelas de nivel primaria por municipio

Municipio	cantidad de escuelas
Álvaro Obregón	181
Azcapotzalco	121
Benito Juárez	133
Coyoacán	157
Cuajimalpa de Morelos	64
Cuauhtémoc	161
Gustavo A. Madero	337
Iztacalco	102
Iztapalapa	378
La Magdalena Contreras	24
Miguel Hidalgo	111
Milpa Alta	27
Tláhuac	58
Tlalpan	162
Venustiano Carranza	125
Xochimilco	93
Total	2234

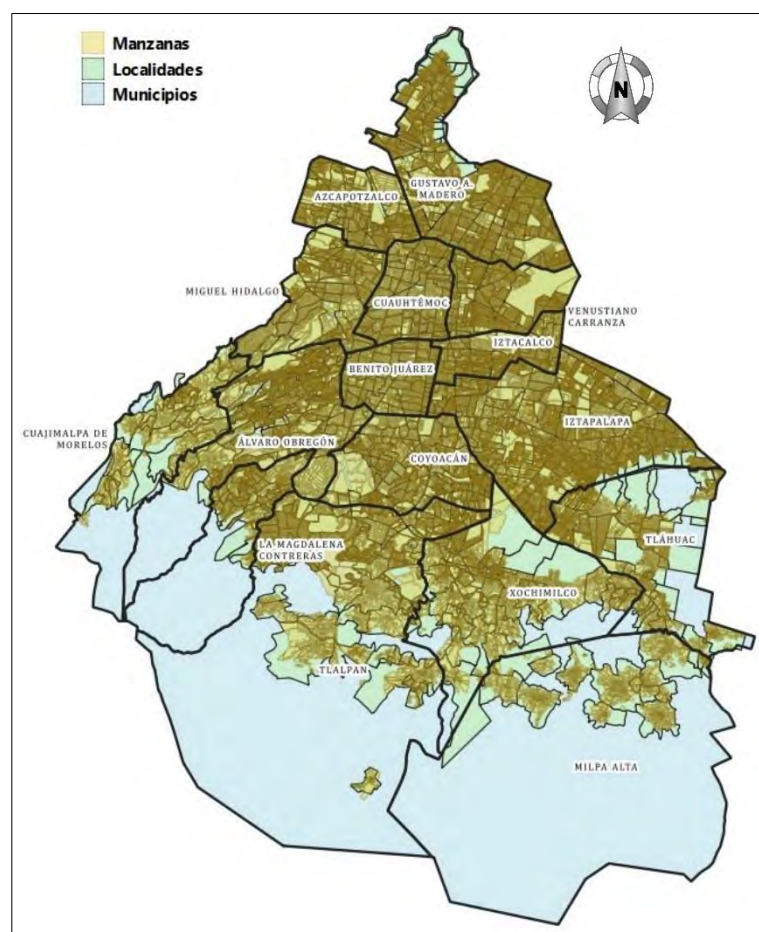


Figura 4.7. Distribución geográfica de la Ciudad de México

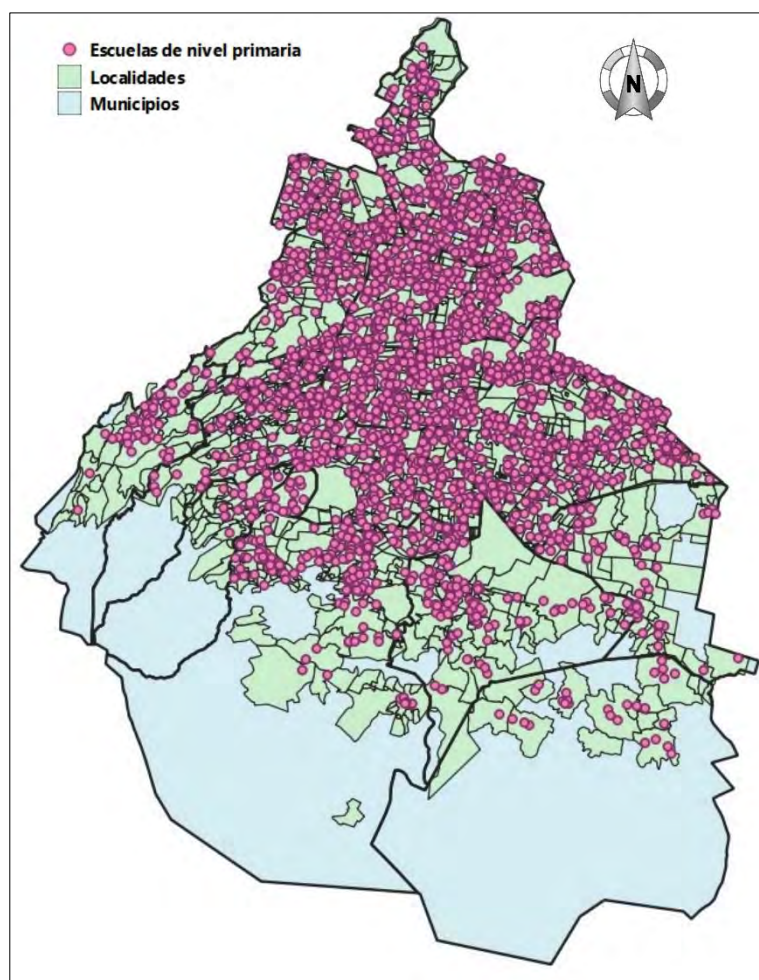


Figura 4.8. Ubicación de las escuelas de nivel primaria

Generación de nuevas variables

Con base en la información previamente recopilada, se procedió a calcular las variables con características espaciales necesarias para completar y enriquecer el conjunto de datos; acorde a la revisión de la literatura. Este proceso incluyó la generación de valores asociados a las relaciones espaciales, garantizando que los datos fueran representativos y adecuados para su posterior análisis.

- Número de escuelas en cada localidad

El número de escuelas se obtuvo utilizando el software QGIS, mediante la función Contar puntos en un polígono, que se encuentra en el menú Vectorial dentro de las Herramientas de Geometría. La Tabla 4.16 presenta ejemplos del número de escuelas por localidad, obtenidos a través del software QGIS.

- o Distancia del trayecto y tiempo de viaje

Posteriormente, se definieron las variables de distancia promedio del trayecto y tiempo promedio de viaje. Ambas variables se calcularon considerando como punto de origen el centroide de cada localidad y como destino la ubicación geográfica de las

escuelas. Tanto la distancia como el tiempo se determinaron utilizando la infraestructura vial y la modalidad de transporte a pie.

El primer paso consistió en calcular la ruta desde el centroide de una localidad hacia cada escuela, expresando la distancia en kilómetros y el tiempo en horas. Para este propósito, se utilizó la plataforma de mapas Geoapify (Geoapify, 2024); la cual ofrece herramientas avanzadas para el desarrollo de mapas, optimización de rutas y análisis geoespacial, así como para la creación de aplicaciones inteligentes basadas en la ubicación geográfica. Geoapify proporciona datos abiertos a través de sus APIs, lo que garantiza información confiable y actualizada. La obtención de las rutas se logró mediante la integración del lenguaje de programación Python, utilizando librerías como *pandas*, y los datos proporcionados por Geoapify.

En la Figura 4.9 se presenta un ejemplo de manera visual del uso de Geoapify, mostrando la ruta desde un punto de origen hacia un destino específico. La Figura 4.10 muestra algunas de las rutas obtenidas desde los centroides de las localidades hacia las escuelas ubicadas en la Ciudad de México.

Finalmente, para calcular los promedios de tiempo y distancia desde el centroide hacia las escuelas en cada localidad, los datos obtenidos se procesaron mediante tablas dinámicas en Microsoft Excel, facilitando el análisis y los cálculos necesarios.

Tabla 4.16. Ejemplos del número de escuelas por localidad

	fid	ENT	CVEDT	NOMDT	NUMPOINTS
1	1 9	2	AZCAPOTZALCO	1	
2	2 9	2	AZCAPOTZALCO	0	
3	3 9	2	AZCAPOTZALCO	3	
4	4 9	2	AZCAPOTZALCO	1	
5	5 9	2	AZCAPOTZALCO	4	
6	6 9	2	AZCAPOTZALCO	7	
7	7 9	2	AZCAPOTZALCO	1	
8	8 9	2	AZCAPOTZALCO	1	

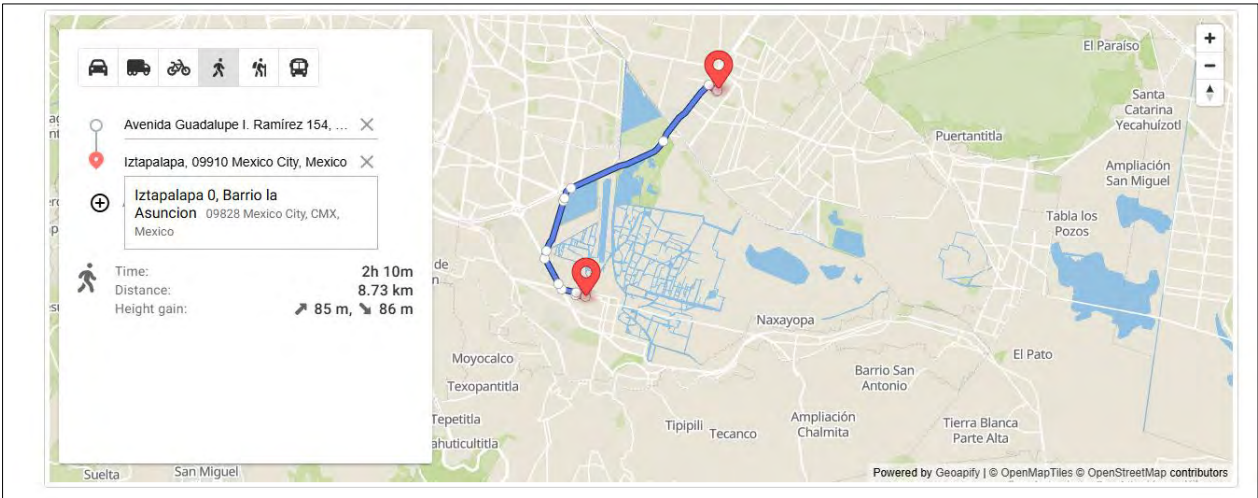


Figura 4.9. Cálculo de la ruta en la plataforma de Geoapify

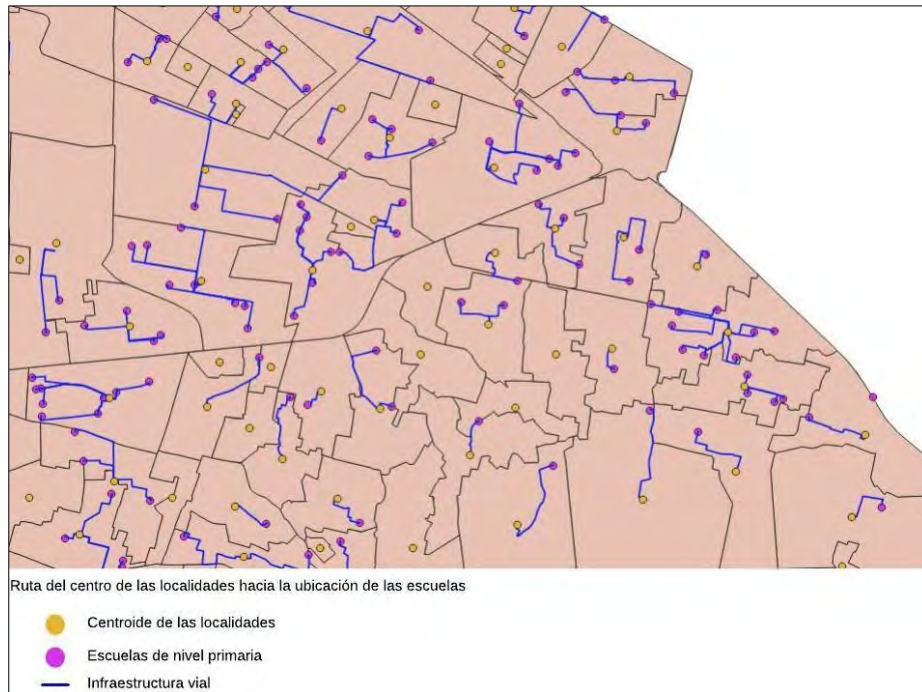


Figura 4.10. Cálculo de las variables de distancia y tiempo

Como resultado, se obtuvo un conjunto de datos que constituye una base sólida e integral para la investigación, al reunir información espacial y contextual relevante para el análisis de la accesibilidad a los conformado por 1,811 instancias, correspondientes a localidades de la Ciudad de México, y 14 variables numéricas, distribuidas de la siguiente manera:

- o 3 variables relacionadas con la información espacial de las localidades.
- o 11 variables correspondientes a aspectos sociales, demográficos y económicos.

La Tabla 4.17 muestra ejemplos de algunas instancias del conjunto, lo que permite visualizar la organización de la información y facilita su interpretación. Adicionalmente, Tabla 4.18 presentan una visión general de la estructura del conjunto de datos integrado, incluyendo el detalle de los atributos considerados. Para el análisis, se dio especial atención a las variables sociales, demográficas y económicas.

Tabla 4.17. Instancias de conjunto de datos para analizar

ID	CVED	NOMDT	CVEUT	NOMUT	Y	X	ANALF	SBASC	OVNAC	PSDS	POBTOT	P6A11	P6A11IN	PORAPR	PPEA	PVPH_SNBIEIN	PVPH_SINTH	NUM_ESC	TIEM_PR_PIE	DIST_PR_PIE
949	7	IZTAPALAPA	07-183	SAN JOSE BU	19.32	-99.034	4.522	32.968	39.633	37.493	5723.000	620.000	3.000	11.495	50.603	0.000	0.000	10.000	10000.000	10000.000
950	7	IZTAPALAPA	07-184	SAN JUAN 2A	19.34	-99.07	5.000	18.579	18.719	37.867	4933.000	419.000	21.000	9.668	55.301	0.000	0.000	1.000	166.799	198.000
951	7	IZTAPALAPA	07-185	SAN JUAN CE	19.34	-99.079	6.099	18.836	22.151	32.518	5984.000	483.000	0.000	9.256	54.880	0.000	0.000	1.000	1520.674	1661.000
952	7	IZTAPALAPA	07-186	SAN JUAN JO	19.34	-99.066	5.000	18.579	18.719	37.867	4847.000	395.000	30.000	9.879	52.094	0.000	0.000	1.000	172.105	204.000
953	7	IZTAPALAPA	07-289	SAN JUAN XA	19.34	-99.083	4.739	17.269	17.961	28.873	13535.000	1004.000	25.000	9.318	56.136	0.074	0.074	0.200	382.359	360.600
954	7	IZTAPALAPA	07-290	SAN JUAN XA	19.33	-99.081	4.739	17.269	17.961	28.873	17917.000	1481.000	47.000	9.388	55.299	0.000	0.061	0.333	551.936	625.667
955	7	IZTAPALAPA	07-188	SAN JUANIC	19.37	-99.119	10.928	15.067	21.183	26.214	6052.000	445.000	3.000	9.924	56.857	0.000	0.328	0.500	129.789	151.000
956	7	IZTAPALAPA	07-189	SAN LORENZ	19.33	-99.074	5.536	28.166	36.077	33.811	1858.000	137.000	3.000	11.820	54.736	0.000	0.000	10.000	10000.000	10000.000
957	7	IZTAPALAPA	07-190	SAN LORENZ	19.3	-99.069	5.136	24.471	31.941	35.112	3742.000	331.000	26.000	9.943	52.111	0.000	0.309	10.000	10000.000	10000.000
958	7	IZTAPALAPA	07-191	SAN LORENZ	19.31	-99.063	3.203	24.824	32.478	35.785	12619.000	1142.000	0.000	10.046	52.159	0.216	0.278	0.500	1062.018	1188.000
959	7	IZTAPALAPA	07-192	SAN LORENZ	19.34	-99.073	4.097	12.894	15.114	27.017	1825.000	87.000	0.000	7.803	56.493	0.000	0.000	10.000	10000.000	10000.000
960	7	IZTAPALAPA	07-193	SAN LORENZ	19.33	-99.068	4.638	12.151	11.969	26.316	1140.000	62.000	40.000	7.970	53.860	0.000	0.000	10.000	10000.000	10000.000
961	7	IZTAPALAPA	07-194	SAN LORENZ	19.38	-99.027	11.461	21.017	24.009	38.344	6301.000	448.000	12.000	9.616	52.674	0.176	0.176	0.200	220.194	236.200
962	7	IZTAPALAPA	07-195	SAN LUCAS	19.36	-99.096	4.675	22.411	31.404	31.724	7582.000	613.000	132.000	8.892	54.300	0.394	0.592	0.200	242.818	273.000
963	7	IZTAPALAPA	07-196	SAN MIGUEL	19.36	-99.078	9.580	16.062	19.400	26.652	28357.000	2118.000	15.000	9.236	55.433	0.000	0.071	0.125	564.841	623.500
964	7	IZTAPALAPA	07-291	SAN MIGUEL	19.35	-98.991	4.638	29.230	37.413	34.387	14947.000	1465.000	32.000	10.660	50.786	0.000	0.000	1.000	208.004	241.000
965	7	IZTAPALAPA	07-292	SAN MIGUEL	19.35	-98.989	3.537	27.162	32.229	35.756	19455.000	2049.000	27.000	11.099	49.787	0.059	0.117	0.111	521.078	516.667
966	7	IZTAPALAPA	07-293	SAN MIGUEL	19.34	-98.987	4.638	29.230	37.413	34.387	16007.000	1661.000	19.000	10.914	51.140	0.070	0.070	0.250	246.584	237.250
967	7	IZTAPALAPA	07-294	SAN MIGUEL	19.33	-98.988	3.537	27.162	32.229	35.756	9955.000	1037.000	6.000	11.045	51.904	0.599	2.247	1.000	632.380	704.000
968	7	IZTAPALAPA	07-198	SAN NICOLAS	19.33	-99.072	3.906	16.144	14.414	31.851	6984.000	504.000	3.000	9.048	54.625	0.000	0.000	0.500	176.998	193.500
969	7	IZTAPALAPA	07-199	SAN NICOLAS	19.33	-99.087	4.112	11.307	10.694	24.260	9413.000	624.000	0.000	8.088	59.482	0.000	0.000	10.000	10000.000	10000.000
970	7	IZTAPALAPA	07-200	SAN PABLO	19.36	-99.089	4.172	20.429	27.136	34.261	5499.000	448.000	28.000	8.752	55.610	0.000	0.206	1.000	215.083	247.000
971	7	IZTAPALAPA	07-201	SAN PABLO	19.33	-99.006	8.113	28.250	38.548	34.078	11880.000	1247.000	14.000	11.161	49.773	0.000	0.000	1.000	902.343	935.000
972	7	IZTAPALAPA	07-202	SAN PEDRO	19.37	-99.085	5.891	24.429	35.542	37.278	6559.000	551.000	15.000	11.021	54.261	0.177	0.236	0.500	1108.322	1209.000

Tabla 4.18. Descripción de la integración del conjunto de datos

Variable	descripción	Tipo de dato
ID	Id de la localidad	-
CVEDT	Clave del municipio	Carácter
NOMDT	Nombre del municipio	Texto
CVEUT	Clave de la localidad	Carácter
NOMUT	Nombre de la localidad	Texto
X	Longitud (ubicación geográfica de la localidad)	Continúo
Y	Latitud (ubicación geográfica de la localidad)	Continúo
ANALF	Porcentaje de la población de 15 años o más analfabeta	Continúo
SBASC	Porcentaje de la población de 15 años o más sin educación básica	Continúo
OVHAC	Porcentaje de viviendas con hacinamiento	Continúo
PSDSS	Porcentaje de la población con acceso a servicios de salud	Continúo
POBTOT	Número de habitantes por localidad	Entero
P6A11	Número de habitantes por localidad de 6 a 11 años de edad	Entero
P6A11NOA	Número de habitantes por localidad de 6 a 11 años de edad que no asiste a la escuela	Entero
PGRAPROES	Grado promedio de escolaridad	Continúo
PPEA	Porcentaje de la población de 12 años y más económicamente activa	Continúo
PVPH_SNBIEN	Porcentaje de viviendas habitadas sin ningún bien	Continúo
PVPH_SINTIC	Porcentaje de viviendas sin tecnologías de la información y de la comunicación (TIC)	Continúo
NUM_ESC	Número de escuelas en la localidad	Entero
TIEM_PR_PIE	Tiempo promedio de viaje del centroide de la localidad hacia la ubicación geográfica de las escuelas en la modalidad a pie.	Continúo (minutos)
DIST_PR_PIE	Distancia promedio del trayecto del centroide de la localidad hacia la ubicación geográfica de las escuelas en la modalidad a pie.	Continúo (metros)

2) Preparación de los datos

La preparación de los datos consiste en la realización de diversas etapas: la imputación de datos faltantes, el cálculo de nuevas variables. Asimismo, la descripción y exploración. Finalmente, la transformación de los datos. Este proceso permite garantizar que los datos sean consistentes y adecuados para el análisis. El objetivo consiste en garantizar que los datos estén limpios, organizados y estructurados para su análisis y desarrollo del modelo matemático de accesibilidad a los servicios educativos mediante la aplicación de diversas técnicas. A continuación, se describen las actividades realizadas en el proceso de recolección de los datos.

Imputación de datos faltantes

En esta actividad, se empleó la técnica de interpolación de vecindad espacial y cálculo del promedio de una zona geográfica (Bautista, 2011). La interpolación de vecindad espacial y cálculo del promedio consiste en identificar las zonas geográficas cercanas a una zona específica; y utilizar los valores conocidos de las áreas vecinas para estimar el valor desconocido de la zona objetivo. Esta técnica permite obtener estimaciones de los diversos variables (variables) al considerar la proximidad espacial de los datos, garantizando que las variaciones en las características geográficas sean reflejadas adecuadamente en el análisis.

Para determinar las localidades vecinas, se utilizó el software de SIG QGIS, por su representación gráfica en un mapa. Mientras que los cálculos matemáticos se llevaron a cabo en Microsoft Excel. El proceso de preparación de datos permitió obtener un conjunto de datos completos, en el cual cada registro cuenta con valores precisos y completos para todos los atributos considerados. Esto garantiza la confiabilidad y representatividad de la información utilizada en el análisis.

La Figura 4.11 ilustra un ejemplo gráfico de la interpolación de vecindad espacial y el cálculo del promedio para estimar el porcentaje de la población sin educación básica en la localidad de Mixcoatl. Este procedimiento se emplea para integrar información de áreas circundantes, mejorando la estimación del valor faltante. De manera similar, este enfoque se aplica para estimar el porcentaje de la población de 12 años y más económicamente activa en la localidad de Industrial II, siguiendo un proceso análogo que garantiza la consistencia y precisión en el análisis de datos, véase Figura 4.12.

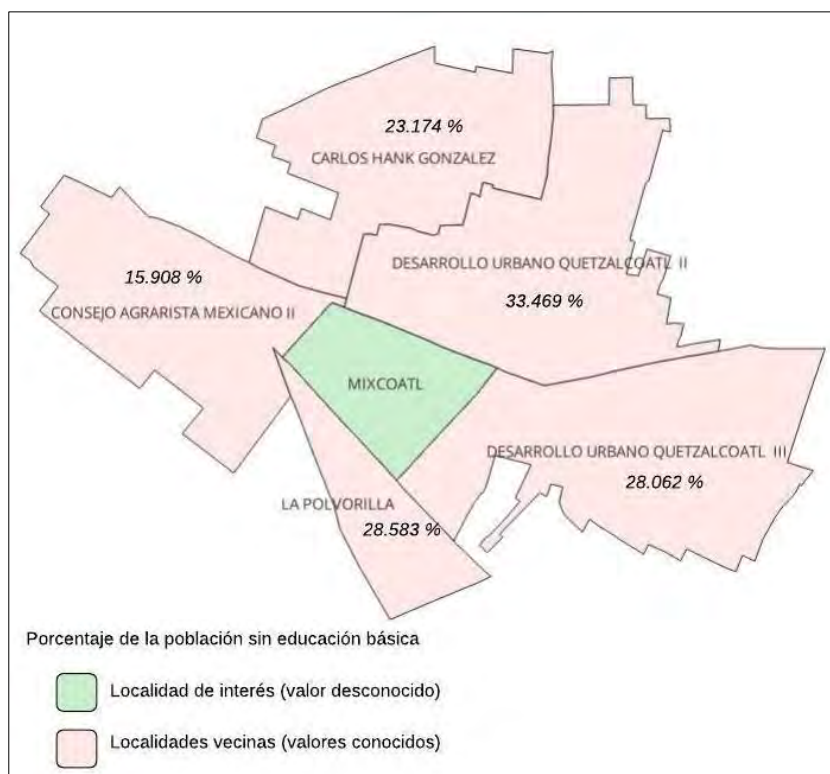


Figura 4.11. Ejemplo 1 de la Interpolación de vecindad espacial

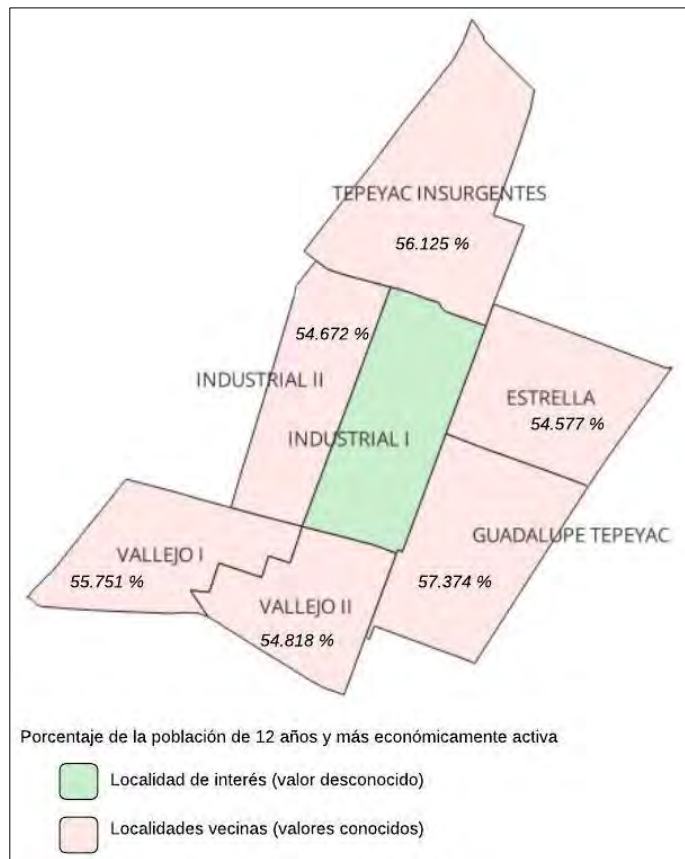


Figura 4.12. Ejemplo 2 de la Interpolación de vecindad espacial

Descripción y exploración de los datos

La estadística se compone de dos áreas la estadística descriptiva y la estadística inferencial. La estadística descriptiva se compone de métodos que son utilizados para describir y sintetizar las características fundamentales de un conjunto de mediciones. Por otro lado, la estadística inferencial abarca métodos diseñados para realizar inferencias basadas en la información obtenida de una muestra representativa de la población. Es decir, permite hacer estimaciones acerca de las características de toda la población a partir de un subconjunto de datos para formular conclusiones (Mendenhall, 2010).

El primer paso consistió en aplicar la estadística descriptiva al conjunto de datos, con el objetivo de resumir y analizar sus principales características, facilitando así su comprensión y exploración inicial. La Tabla 4.19 y la Tabla 4.20 muestra la descripción del conjunto de datos; donde se puede observar que el conjunto contempla variables numéricas de diversas escalas y existen registros faltantes. Los cálculos se realizaron utilizando el lenguaje R, empleando la librería *pastecs*.

Tabla 4.19. Descripción del conjunto de datos, parte 1

	FID	CVEDT	NOMDT	CVEUT	NOMUT	Y	X	ANALF	SBASC	OVHAC	PSDSS	POBTOT
Total de registros	1811	1811	-	-	-	1811	1811	1811	1811	1811	1811	1811
Valores nulos	0	0	-	-	-	0	0	0	0	1	0	0
Datos inconsistentes	0	0	-	-	-	0	0	0	0	0	0	0
Mínimo	1	2	-	-	-	19.135	-99.342	1	1	0	0.030	2.163
Máximo	1814	17	-	-	-	19.581	-98.960	5	5	5	25.142	44.921
Rango	1813	15	-	-	-	0.447	0.383	4	4	5	25.111	42.758
Suma	1644056	15847	-	-	-	35085.671	-179547.364	3812.000	3793.000	5109.000	9712.863	30759.869
Mediana	908	8	-	-	-	19.365	-99.141	2	2	3	4.963	17.479
Media	907.8167	8.750414	-	-	-	19.374	-99.143	2.105	2.094	2.821	5.363	16.985
SE.mean	12.30908	0.10182	-	-	-	0.002	0.002	0.021	0.032	0.043	0.059	0.170
CI.mean.0.95	24.1415	0.199697	-	-	-	0.004	0.003	0.041	0.062	0.084	0.117	0.334
Varianza	274391	18.77524	-	-	-	0.006	0.005	0.798	1.826	3.311	6.406	52.426
Desviación estándar	523.8234	4.333041	-	-	-	0.080	0.069	0.893	1.351	1.819	2.531	7.241
Coefficiente de variación	0.577015	0.495181	-	-	-	0.004	-0.001	0.424	0.645	0.645	0.472	0.426

Tabla 4.20. Descripción del conjunto de datos, parte 2

	P6A11	P6A11NOA	PGRAPROES	PPEA	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	TIEM_PR_PIE	DIST_PR_PIE
Total de registros	1811	1811	1811	1811	1811	1811	1811	1811	1811
Valores nulos	0	0	0	0	0	0	0	0	1
Datos inconsistentes	0	0	0	0	0	0	0	0	0
Mínimo	3	0	6.221	17.443	0.000	0.000	0.067	0.050	0
Máximo	3491	132	31.250	76.188	2.421	5.306	10	10000	10000
Rango	3488	132	25.029	58.745	2.421	5.306	9.933	9999.950	10000
Suma	805920.000	18495.000	16421.898	99875.931	63.688	163.516	9315.824	9155933.352	9184133.841
Mediana	309	5	9.061	54.748	0	0	1	1325.264	1275.000
Media	445.014	10.213	9.068	55.150	0.035	0.090	5.144	5055.733	5071.305
SE.mean	10.072	0.351	0.034	0.104	0.003	0.007	0.110	112.272	111.952
CI.mean.0.95	19.753	0.689	0.067	0.203	0.007	0.013	0.216	220.197	219.568
Varianza	183699.141	223.575	2.123	19.405	0.021	0.080	21.969	22827790.057	22697554.517
Desviación estándar	428.601	14.952	1.457	4.405	0.146	0.282	4.687	4777.844	4764.195
Coefficiente de variación	0.963	1.464	0.161	0.080	4.148	3.124	0.911	0.945	0.939

Para lograr una mayor comprensión de los datos, se emplearon técnicas estadísticas gráficas como histogramas y diagramas de caja, para visualizar las distribuciones y el comportamiento de las variables. Las Figura 4.13, Figura 4.14 y Figura 4.15 presentan las gráficas de visualización del histograma de frecuencias de cada criterio. Asimismo, Figura 4.16, Figura 4.17 y Figura 4.18 muestran los diagramas de cajas para cada uno de las variables. Los gráficos se generaron utilizando el lenguaje R, empleando las funciones base hist y boxplot. Por otro lado, las figuras, muestran el mapa de coropletas de la representación en los mapas para cada uno de las variables. Estas herramientas gráficas facilitaron el análisis detallado de la forma, la dispersión y los posibles valores atípicos de los datos.

Adicionalmente, las variables fueron representadas en un mapa mediante técnicas de estadística geoespacial, con el propósito de analizar e interpretar la distribución espacial de los datos y que permitan proporcionar una comprensión más profunda de su comportamiento en relación con la geografía del área de estudio. Para este análisis, se utilizó el software QGIS, a través de la herramienta de propiedades, se aplicó una simbología graduada para generar mapas de coropletas. Un mapa de coropletas es un tipo de mapa temático que representa valores numéricos de zonas geográficas

mediante colores o sombras. Esta visualización es útil para analizar la distribución espacial de variables y detectar patrones o variaciones dentro de una región. La representación de los datos depende del método de clasificación, que agrupa los valores en intervalos para facilitar su interpretación y comparación (Ribecca, 2025), (CURSOSONLINEGIS, 2025).

Se empleó el método de clasificación de Jenks, también conocido como Natural Breaks, el cual agrupa los valores de manera óptima dividiendo los datos en intervalos significativos que reflejan mejor su estructura. Este método minimiza la varianza dentro de cada grupo y maximiza la diferencia entre ellos, identificando patrones en la distribución de los datos. Su objetivo es generar clases que resalten las diferencias relevantes, facilitando así la interpretación visual en mapas o gráficos. La Figura 4.19, Figura 4.20, Figura 4.21, Figura 4.22, Figura 4.23, Figura 4.24 y la Figura 4.25 presentan el mapa de coropletas, los valores de la clasificación por intervalos y la gráfica de comportamiento basada en dicha clasificación.

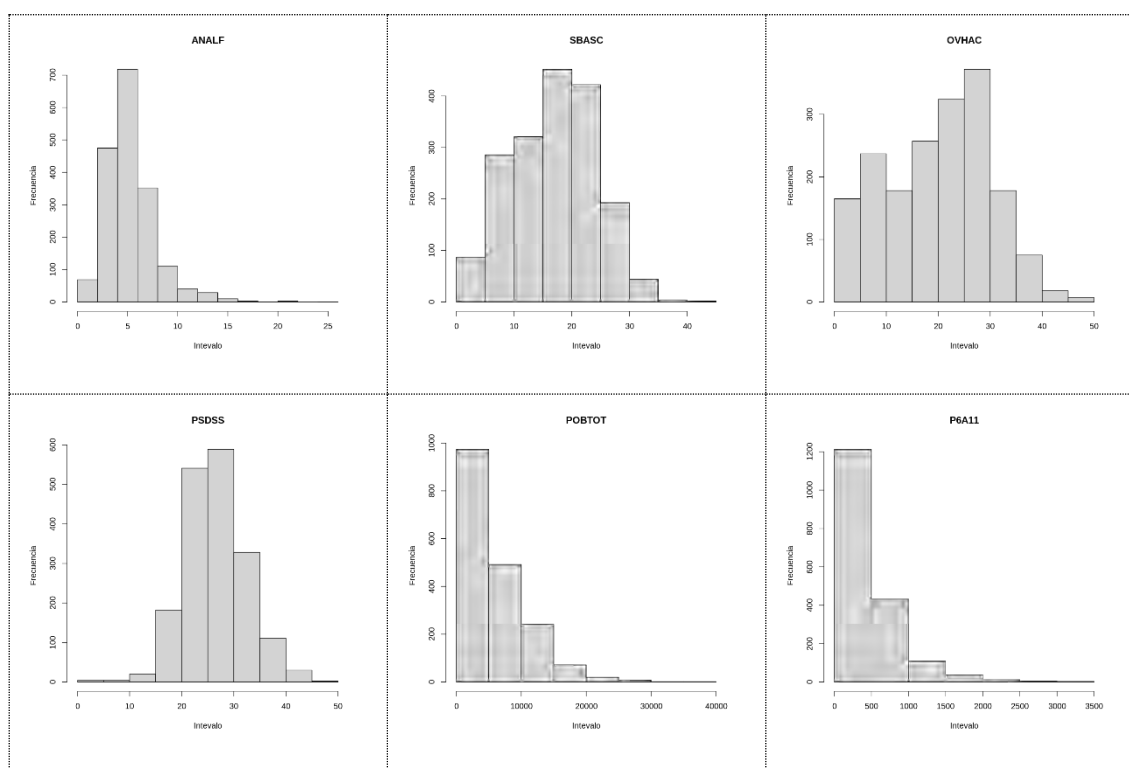


Figura 4.13. Histogramas de frecuencia de las variables. Parte 1

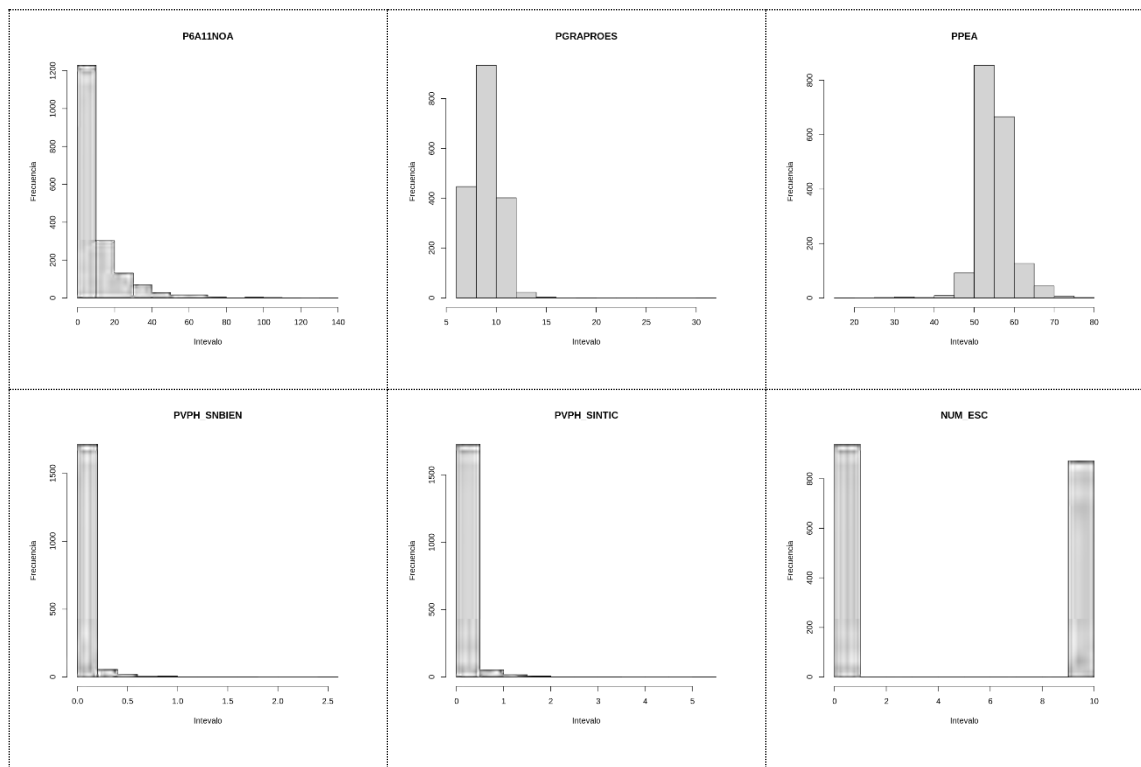


Figura 4.14. Histogramas de frecuencia de las variables. Parte 2

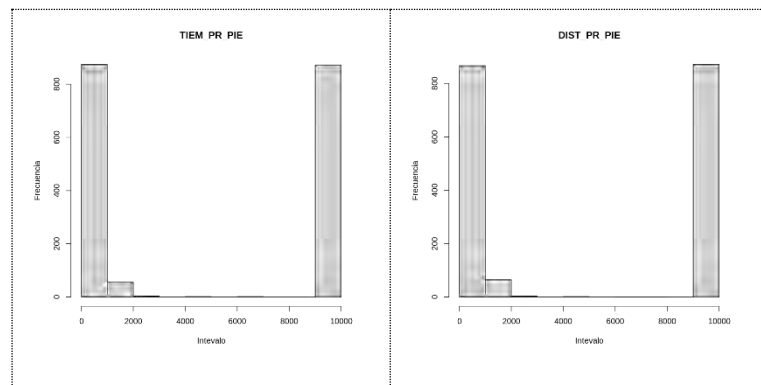


Figura 4.15. Histogramas de frecuencia de las variables. Parte 3

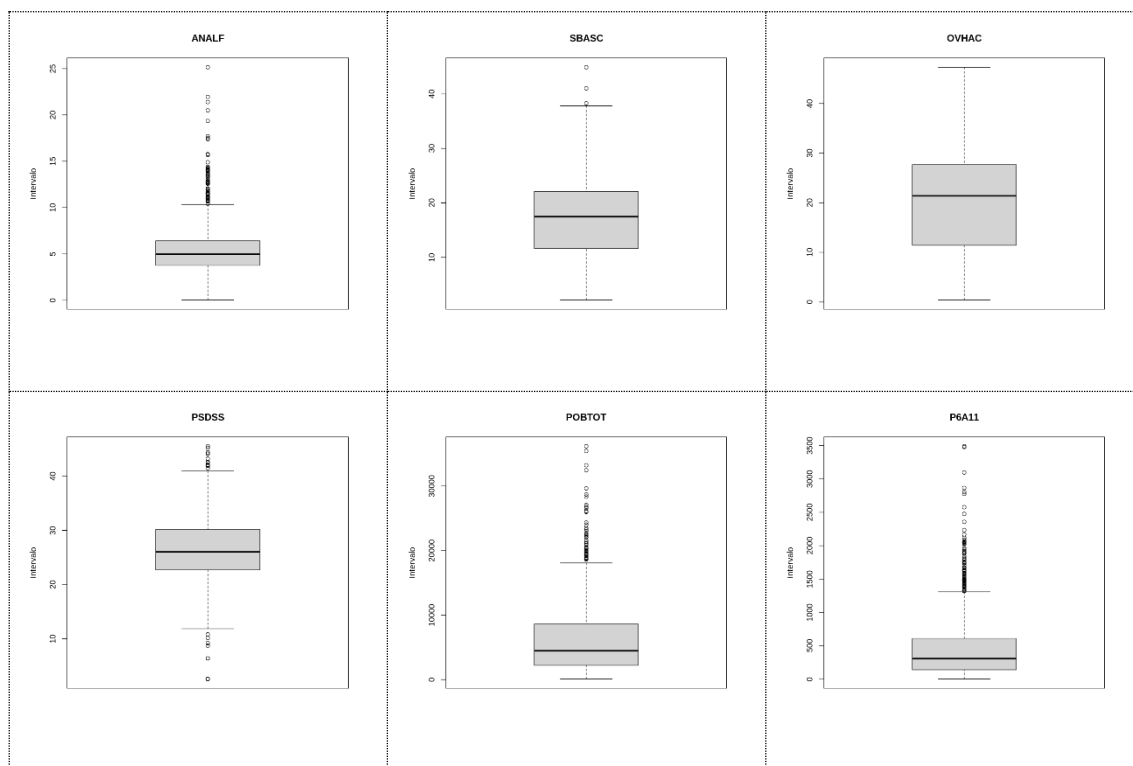


Figura 4.16. Diagramas de caja de las variables. Parte 1

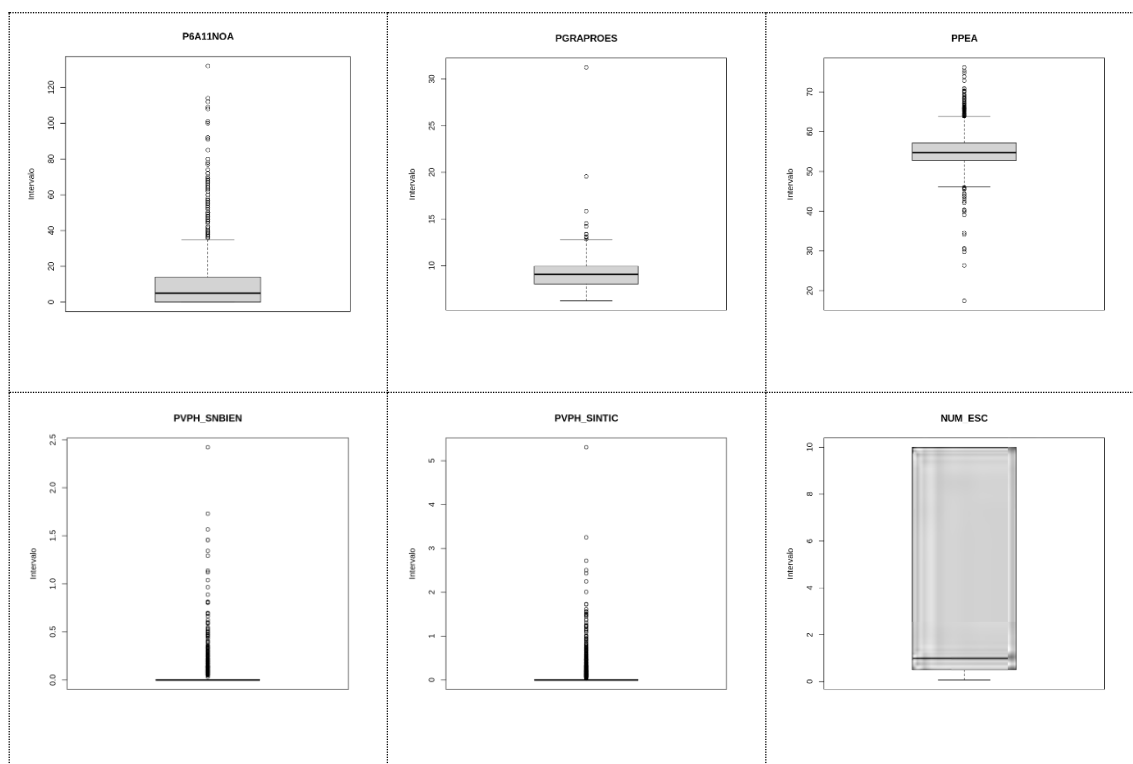


Figura 4.17. Diagramas de caja de las variables. Parte 2

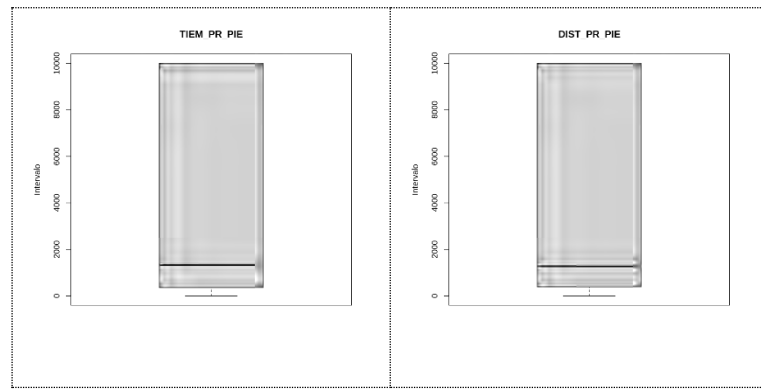


Figura 4.18. Diagramas de caja de las variables. Parte 3

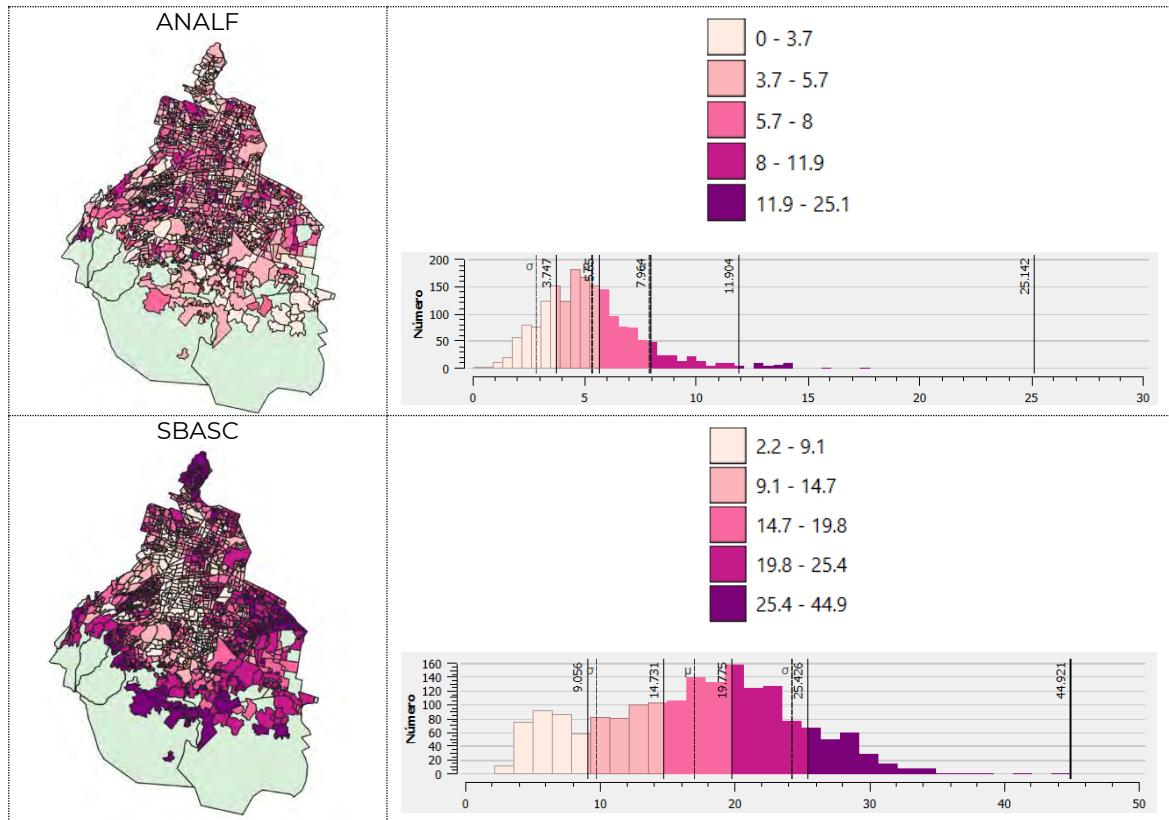


Figura 4.19. Mapa de coropletas de las variables Parte 1

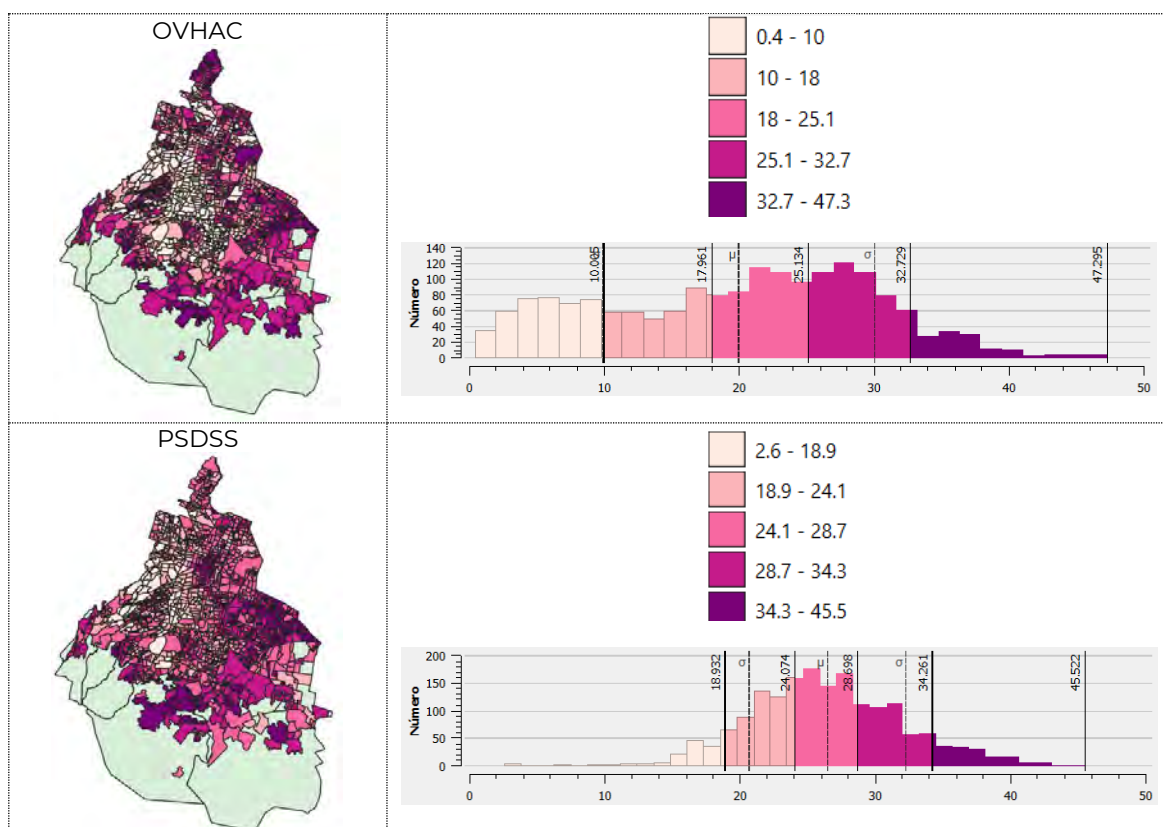


Figura 4.20. Mapa de coropletas de las variables. Parte 2

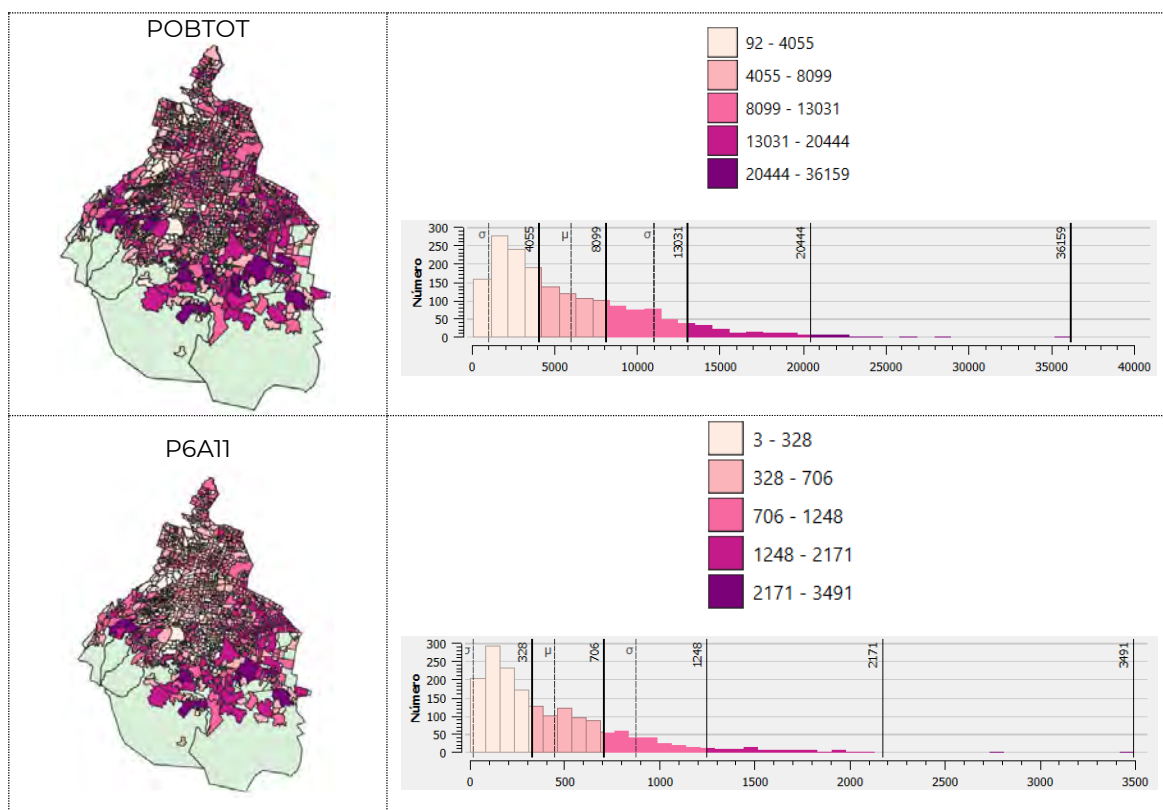


Figura 4.21. Mapa de coropletas de las variables. Parte 3

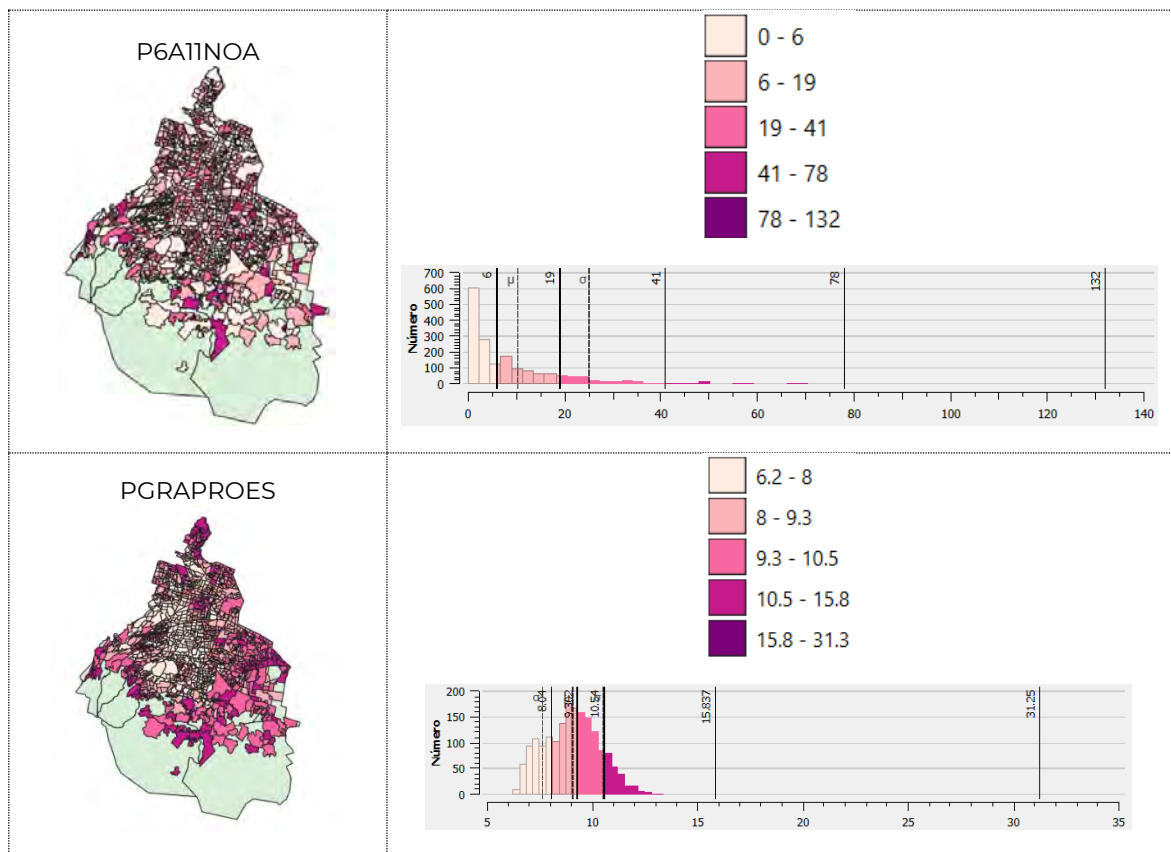


Figura 4.22. Mapa de coropletas de las variables. Parte 4

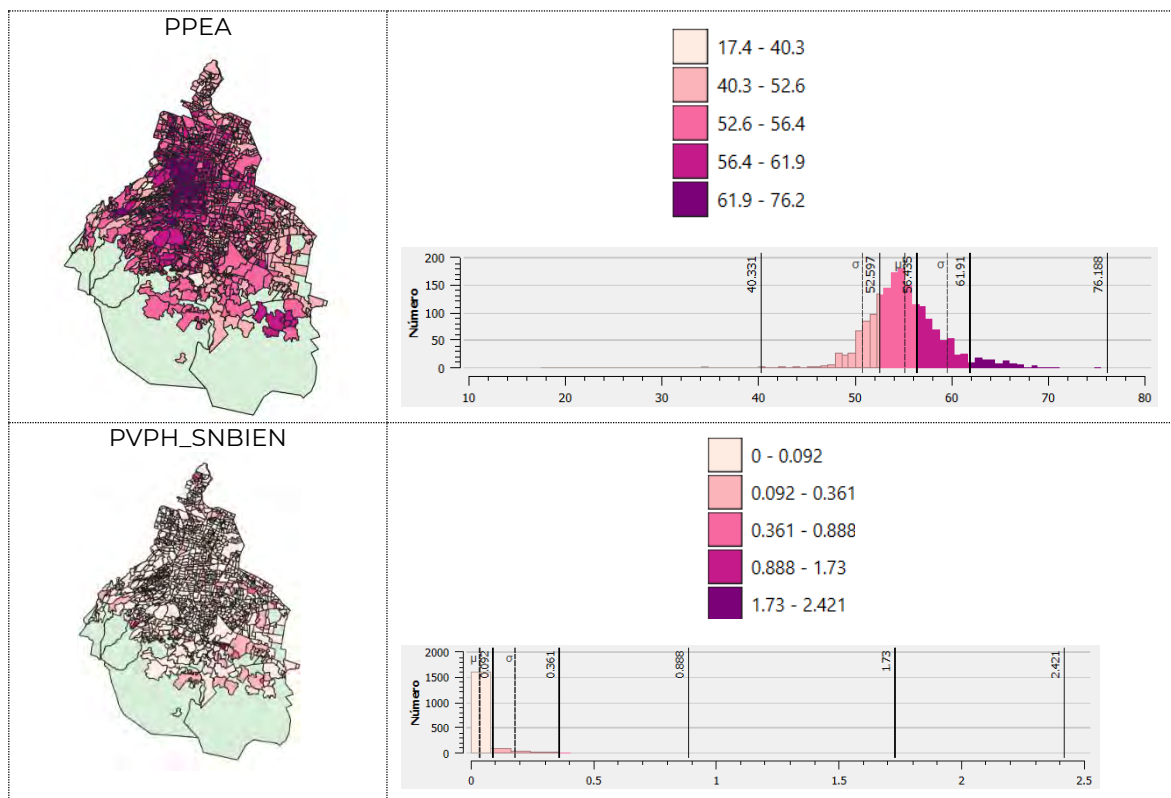


Figura 4.23. Mapa de coropletas de las variables. Parte 5

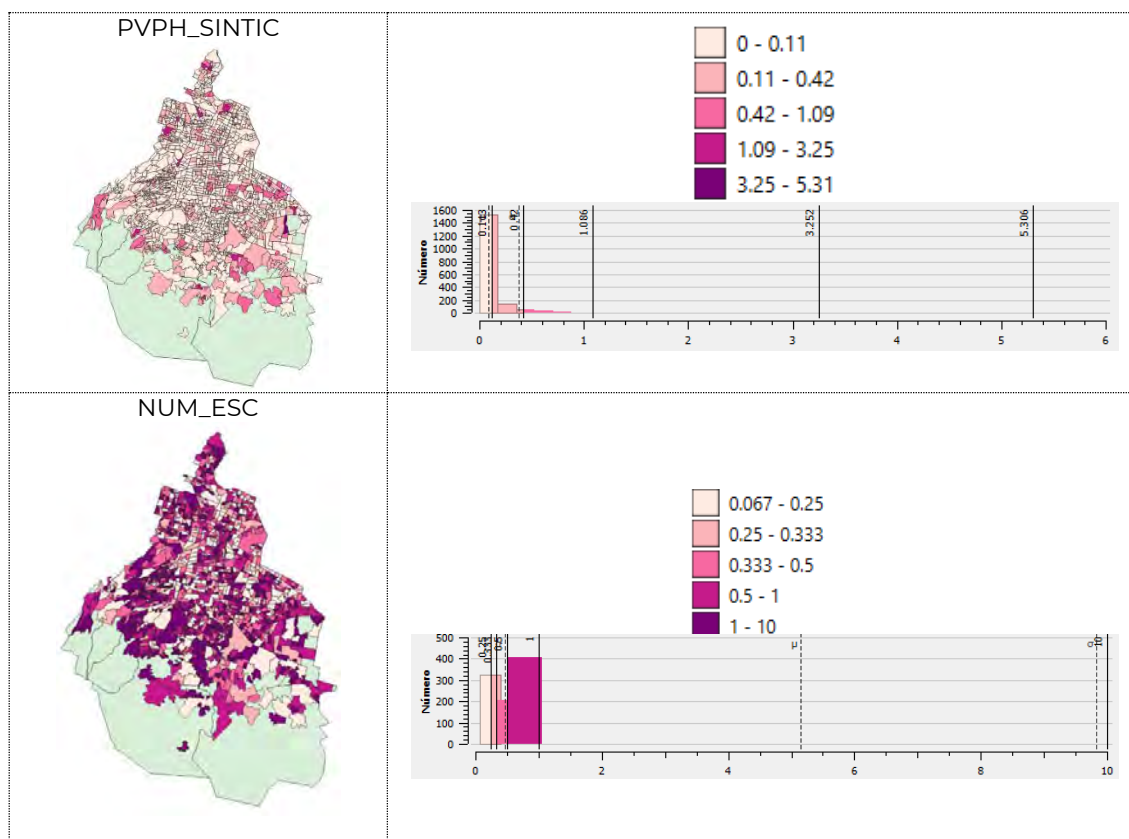


Figura 4.24. Mapa de coropletas de las variables. Parte 6

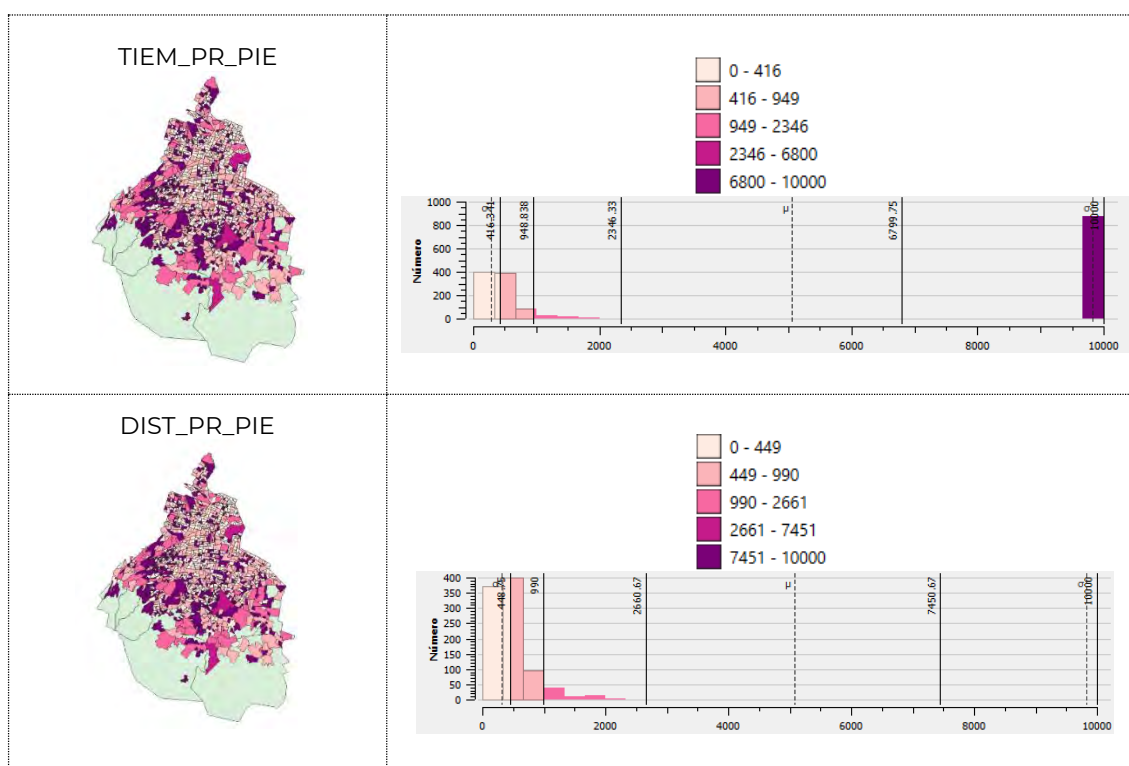


Figura 4.25. Mapa de coropletas de las variables. Parte 7

De este análisis, se pueden extraer algunas conclusiones principales sobre el conjunto de datos.

- El conjunto de datos cuenta con 1,811 valores válidos y completos.
- Las variables P6A11NOA (población de 6 a 11 años que no acude a la escuela) y PVPH_SNBIEN (población sin acceso algún bien) tienen coeficientes de variación grandes, lo que implica una alta variabilidad relativa con respecto a la media.
- Los valores de la media y la mediana de las variables son cercanos. Esto indica una distribución relativamente simétrica. Sin embargo, hay excepciones POBTOT (población total) y DIST_PR_PIE (distancia de la localidad hacia las escuelas caminando), donde existen diferencias.
- Las variables de POBTOT, DIST_PR_PIE y TIEM_PR_PIE (tiempo de viaje de la localidad hacia las escuelas caminando) tienen desviaciones estándar altas, lo que indica que los datos están muy dispersos.
- Destaca la diversidad entre las instancias, con respecto a los valores mínimos y máximos de las variables POBTOT y DIST_PR_PIE.
- En algunas variables se pueden observar patrones de comportamiento espacial. Por ejemplo, PGRAPROES (grado de escolaridad) y SBASC (población sin educación básica) sus valores más pequeños tienen a ubicarse hacia el centro de la ciudad.
- En los diagramas de caja, se observan valores atípicos. Por ejemplo, P6A11 (población con edad de 6 a 11 años).

En el marco de la estadística inferencial, en esta etapa, se realizaron pruebas de normalidad y densidad con el objetivo de analizar la estructura y el comportamiento de las variables con respecto a la distribución normal. Como primer paso, se aplicó la prueba de normalidad de Shapiro-Wilk (Pulido, H.; De la Vara, 2008), cuyo planteamiento de hipótesis se presenta en la ecuación (5.1), considerando un nivel de significancia del 5%. Los resultados obtenidos, presentados en las Tabla 4.21 y Tabla 4.22 indicaron que las variables no siguen una distribución normal. Posteriormente, estos hallazgos fueron corroborados de manera visual mediante las gráficas de densidad de la Figura 4.26, Figura 4.27, Figura 4.28; y las gráficas de normalidad en la Figura 4.29, Figura 4.30, Figura 4.31.

Planteamiento de la prueba de hipótesis para la prueba de normalidad Shapiro-Wilks

(4.17)

H_0 : El criterio sigue una distribución normal

H_1 : El criterio no sigue una distribución normal

($\alpha = 0.05$)

Tabla 4.21. Resultados de la prueba Shapiro-Wilks de algunas variables. Parte 1

Shapiro-Wilk normality test data: Data\$ANALF W = 0.88659, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$SBASC W = 0.98487, p-value = 6.743e-13	Shapiro-Wilk normality test data: Data\$OVHAC W = 0.97446, p-value < 2.2e-16
Shapiro-Wilk normality test data: Data\$PSDSS W = 0.99214, p-value = 2.815e-08	Shapiro-Wilk normality test data: Data\$POBTOT W = 0.86401, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$P6A11 W = 0.80466, p-value < 2.2e-16
Shapiro-Wilk normality test data: Data\$P6A11NOA W = 0.68428, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$PGRAPROES W = 0.89872, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$PPEA W = 0.91744, p-value < 2.2e-16

Tabla 4.22. Resultados de la prueba Shapiro-Wilks de algunas variables. Parte 2

Shapiro-Wilk normality test data: Data\$PVPH_SNBIEN W = 0.2562, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$PVPH_SINTIC W = 0.34778, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$NUM_ESC W = 0.66716, p-value < 2.2e-16
Shapiro-Wilk normality test data: Data\$TIEM_PR_PIE W = 0.66205, p-value < 2.2e-16	Shapiro-Wilk normality test data: Data\$DIST_PR_PIE W = 0.66326, p-value < 2.2e-16	

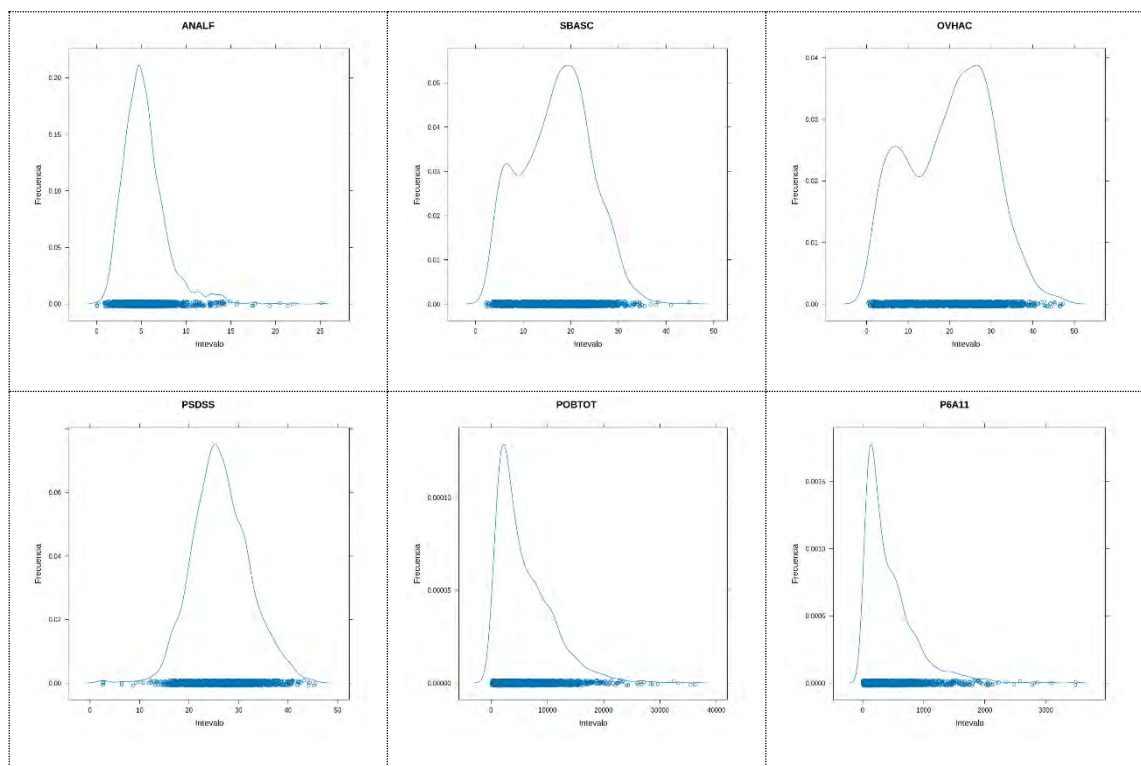


Figura 4.26. Resultados de la prueba de densidad. Parte 1

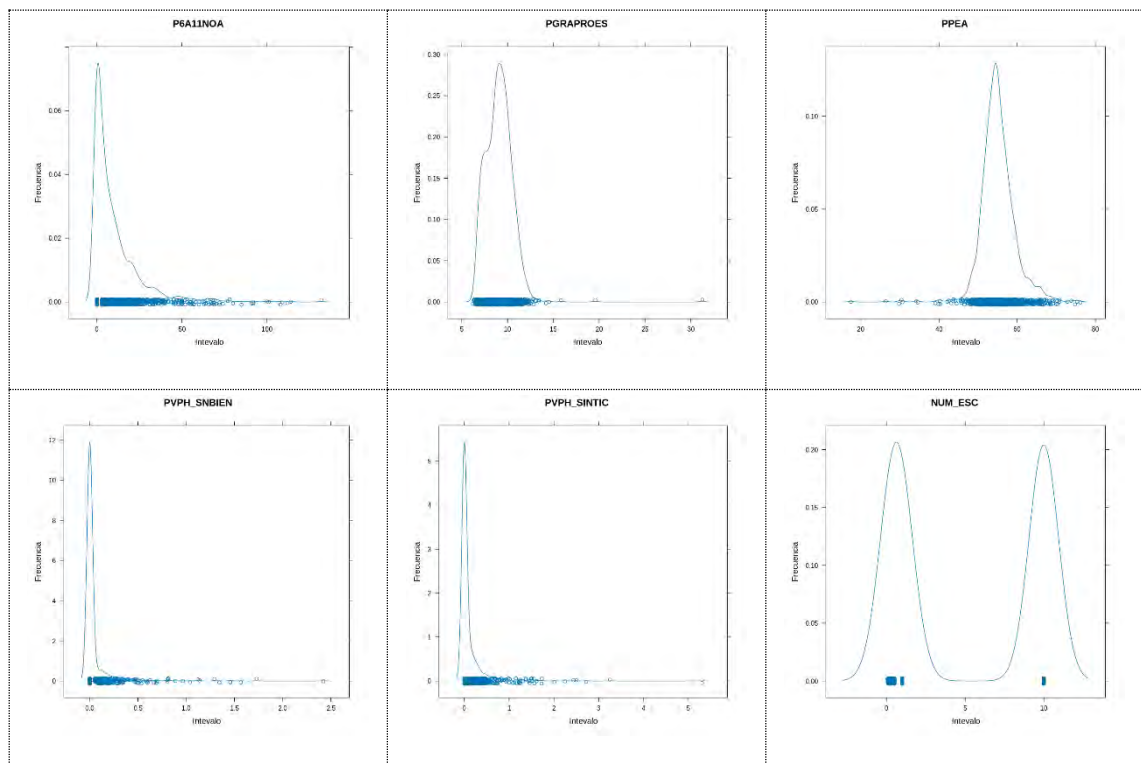


Figura 4.27. Resultados de la prueba de densidad. Parte 2

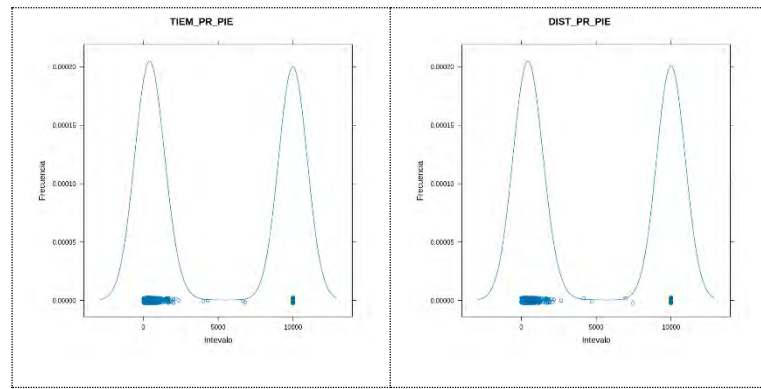


Figura 4.28. Resultados de la prueba de densidad. Parte 3

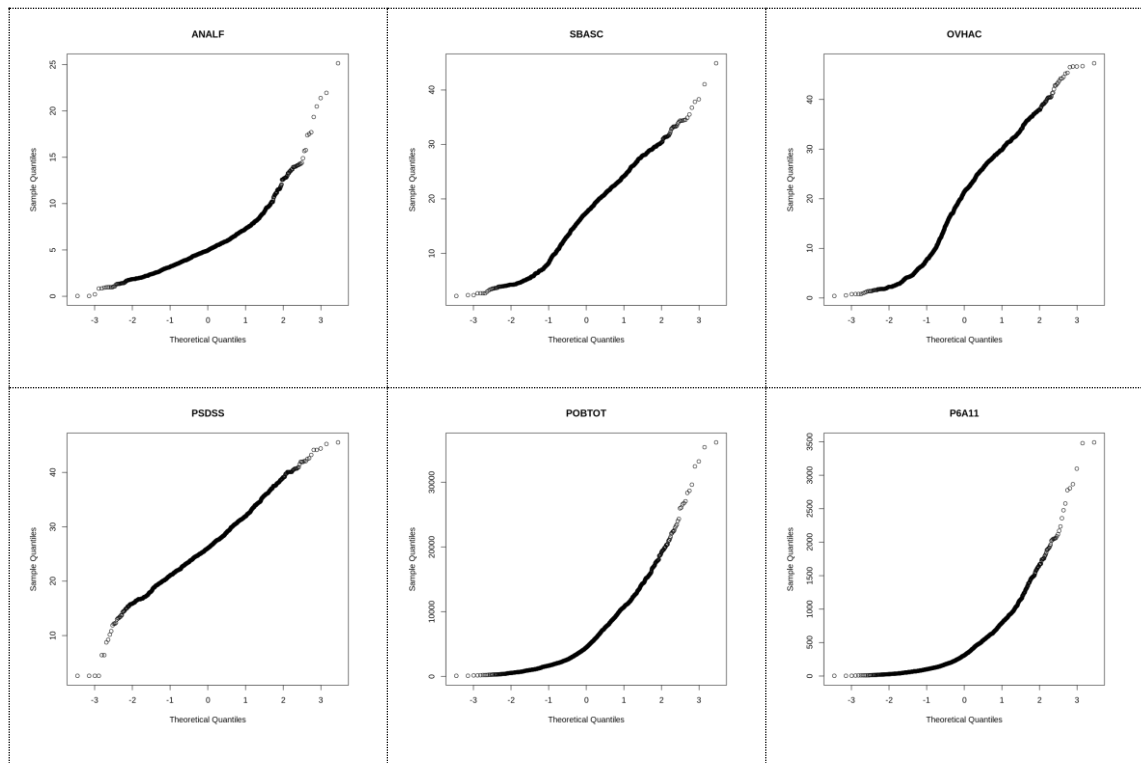


Figura 4.29. Resultados de la prueba de normalidad. Parte 1

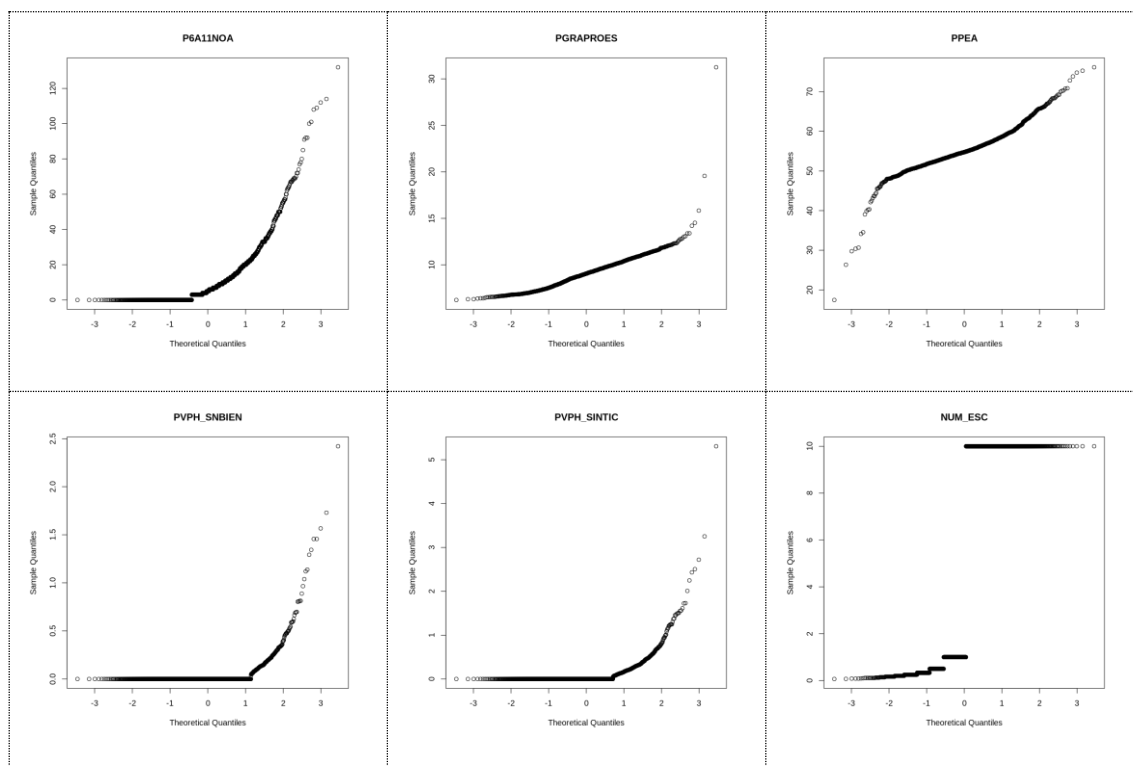


Figura 4.30. Resultados de la prueba de normalidad. Parte 2

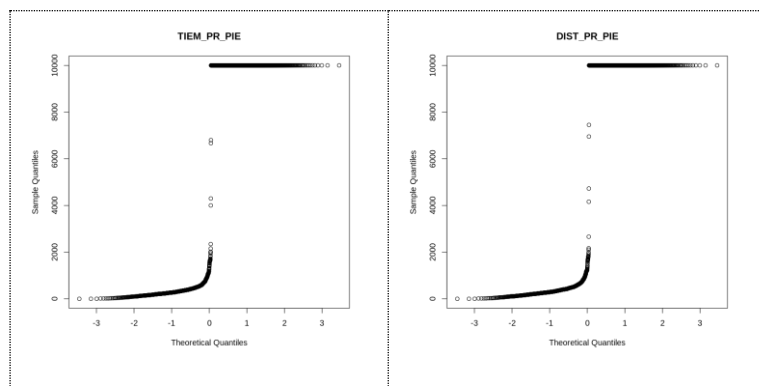


Figura 4.31. Resultados de la prueba de normalidad. Parte 3

Tratamiento de los datos

En esta etapa, se llevó a cabo el tratamiento de algunas variables del conjunto de datos con el objetivo de garantizar la comparabilidad entre las variables analizadas. Para ello, se aplicó una transformación recíproca, la cual convierte un valor x en su inverso multiplicativo $\frac{1}{x}$. Esta transformación suele ser útil cuando se requiere interpretar valores más altos como menos favorables y valores más bajos como más favorables, o viceversa.

En particular, para el criterio NUM_ESC, que representa el número de escuelas, se implementó la transformación recíproca con el propósito de reflejar la accesibilidad de manera más intuitiva. De este modo, un mayor número de escuelas se asocia con una mayor accesibilidad, mientras que valores más altos después de la transformación reflejan una menor accesibilidad, véase ecuación 4.2.

Asimismo, se realizaron ajustes adicionales en las variables TIEM_PR_PIE y DIST_PR_PIE, con el fin de mejorar su interpretación y coherencia dentro del análisis, como se muestra en la ecuación 4.18-4.20.

La razón detrás de esta transformación a las variables, radica en la necesidad de mantener la coherencia entre todas las variables del conjunto de datos. En dicho conjunto, los valores elevados reflejan localidades en condiciones menos favorables; mientras que los valores más bajos representan localidades con mejores condiciones.

$$f(x) = \begin{cases} \frac{1}{NUM_ESC} & \text{si } NUM_ESC \neq 0 \\ 10 & \text{si } NUM_ESC = 0 \end{cases} \quad (4.18)$$

$$g(x) = \begin{cases} TIEM_PR_PIE & \text{si } TIEM_PR_PIE \neq 0 \\ 10000 & \text{si } TIEM_PR_PIE = 0 \end{cases} \quad (4.19)$$

$$h(x) = \begin{cases} DIST_PR_PIE & \text{si } DIST_PR_PIE \neq 0 \\ 10000 & \text{si } DIST_PR_PIE = 0 \end{cases} \quad (4.20)$$

Transformación de los datos

La transformación de datos es un paso fundamental cuando las variables presentan diferentes escalas; ya que permite su comparabilidad y un análisis más preciso. En este conjunto de datos, las variables se encuentran en distintas escalas, por lo que se aplicó la técnica de normalización Z-score. Esta técnica transforma los valores originales para que tengan una media de cero y una desviación estándar de uno, facilitando así su interpretación y uso en análisis posteriores. La ecuación 4.5 presenta la fórmula matemática utilizada para el cálculo del Z-score.

El proceso de transformación se llevó a cabo utilizando el lenguaje de programación R, específicamente la función `scale` del paquete `base`. Asimismo, se realizó un análisis descriptivo de los datos normalizados, donde se pueden observar la media y la desviación estándar, como se muestra en la Tabla 4.23 y Tabla 4.24.

$$z = \frac{x - \mu}{\sigma} \quad (4.21)$$

Donde:

z : es el valor transformado.

x es el valor original de los datos.

μ es la media de los datos de la variable.

σ es la desviación estándar de los datos de la variable.

Tabla 4.23. Descripción del conjunto de datos transformado, parte 1

	ANALF	SBASC	OVHAC	PSDSS	POBTOT	P6ATI	P6ATIIOA
Total de registros	1811	1811	1811	1811	1811	1811	1811
Valores nulos	0	0	0	0	0	0	0
Datos inconsistentes	0	0	0	0	0	0	0
Mínimo	-2.107	-2.047	-1.945	-4.110	-1.184	-1.031	-0.683
Máximo	7.814	3.858	2.718	3.289	6.039	7.107	8.145
Rango	9.921	5.905	4.663	7.399	7.223	8.138	8.828
Suma	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mediana	-0.158	0.068	0.141	-0.067	-0.309	-0.317	-0.349
Media	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SE.mean	0.023	0.023	0.023	0.023	0.023	0.023	0.023
CI.mean.0.95	0.046	0.046	0.046	0.046	0.046	0.046	0.046
Varianza	1	1	1	1	1	1	1
Desviación estándar	1	1	1	1	1	1	1
Coefficiente de variación	-	-	-	-	-	-	-
	6.8E+15	1.2E+16	6.3E+15	5.0E+15	1.4E+16	3.0E+16	1.5E+16

Tabla 4.24. Descripción del conjunto de datos transformado, parte 2

	PGRAPROES	PPEA	PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	TIEM_PR_PIE	DIST_PR_PIE
Total de registros	1811	1811	1811	1811	1811	1811	1811
Valores nulos	0	0	0	0	0	0	0
Datos inconsistentes	0	0	0	0	0	0	0
Mínimo	-1.954	-8.560	-0.241	-0.320	-1.083	-1.058	-1.064
Máximo	15.224	4.776	16.355	18.493	1.036	1.035	1.035
Rango	17.178	13.336	16.596	18.813	2.119	2.093	2.099
Suma	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mediana	-0.005	-0.091	-0.241	-0.320	-0.884	-0.781	-0.797
Media	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SE.mean	0.023	0.023	0.023	0.023	0.023	0.023	0.023
CI.mean.0.95	0.046	0.046	0.046	0.046	0.046	0.046	0.046
Varianza	1	1	1	1	1	1	1
Desviación estándar	1	1	1	1	1	1	1
Coefficiente de variación	2.9E+15	-2.5E+15	1.1E+18	-1.3E+17	-1.2E+16	4.1E+16	5.9E+16

Todas las actividades relacionadas el preprocesamiento de los datos fueron esenciales para obtener un conjunto de datos adecuado en el contexto de la accesibilidad a los servicios (escuelas primarias) y listo para su análisis.

El conjunto de datos final, preparado para el análisis, está conformado por 1,811 instancias correspondientes a diversas localidades. Incluye 12 variables exploratorias que reflejan características sociales, demográficas, económicas y espaciales. Tabla 4.25 muestra un subconjunto de los datos transformados.

Tabla 4.25. Subconjunto de datos transformados

FID	CVEDT	NOMD	CVEU	NOMU	Y	X	ANALF	SBASC	OVHAC	PSDSS	POSTOT	PA611	PA611NOA	PGRAPROES	PPEA	PVPH	SNBIEN	PVPH_SINTIC	NUM_ESC	TIEM	PR_PIE	DIST	PR_PIE
691	5	GUSTAVO 05-197	VERONICA	19.573	-99.129	-0.859	2.226	2.660	1.000	-0.724	-0.520	-0.683	1.850	-1.592	-0.241	-0.320	1.036	1.035	1.035				
692	5	GUSTAVO 05-213	VILLA DE J	19.464	-99.058	0.558	-0.732	-1.101	0.055	-1.049	-0.901	-0.148	1.567	-0.172	-0.241	-0.320	1.036	1.035	1.035				
693	5	GUSTAVO 05-198	VILLA DE J	19.468	-99.061	0.558	-0.732	-1.101	0.055	-0.229	-0.329	-0.482	0.383	-1.127	-0.241	-0.320	-0.991	-1.030	-1.031				
694	5	GUSTAVO 05-199	VILLA GU	19.484	-99.114	-0.366	0.036	0.564	0.202	-0.429	-0.273	-0.282	-0.077	0.548	-0.241	-0.320	-1.026	-0.995	-0.996				
695	5	GUSTAVO 05-200	VILLA HER	19.502	-99.092	3.468	0.394	0.535	0.728	-0.701	-0.595	-0.683	0.183	-0.756	-0.241	-0.320	-0.884	-1.054	-1.059				
696	5	GUSTAVO 05-202	VISTA HEF	19.548	-99.123	-0.765	2.256	2.367	2.448	-0.709	-0.413	-0.683	2.069	-1.012	-0.241	-0.320	1.036	1.035	1.035				
697	5	GUSTAVO 05-251	ZONA ESC	19.539	-99.150	-0.936	0.589	0.178	0.292	0.224	0.289	-0.482	0.129	-0.882	-0.241	-0.320	-0.884	-1.015	-1.019				
698	5	GUSTAVO 05-252	ZONA ESC	19.542	-99.147	-0.936	0.589	0.178	0.292	0.243	0.231	-0.683	0.254	-0.678	-0.241	-0.320	1.036	1.035	1.035				
699	5	GUSTAVO 05-205	ZONA ESC	19.540	-99.143	-1.015	0.149	-0.388	-0.435	-0.564	-0.455	0.120	0.240	-0.382	-0.241	-0.320	-0.884	-1.037	-1.042				
700	6	IZTACALCI 06-040	AGRICOLA	19.402	-99.077	-0.901	-0.190	-0.094	0.323	2.229	1.864	-0.081	-0.171	0.322	-0.241	-0.320	-0.991	-0.901	-0.890				
701	6	IZTACALCI 06-041	AGRICOLA	19.396	-99.077	-0.205	-0.328	-0.301	-0.240	0.831	0.763	0.053	0.108	0.087	-0.241	0.132	-1.026	-0.945	-0.942				
702	6	IZTACALCI 06-042	AGRICOLA	19.390	-99.076	-0.901	-0.190	-0.094	0.323	1.132	0.782	-0.482	-0.236	-0.468	-0.241	-0.320	-1.055	-0.937	-0.932				
703	6	IZTACALCI 06-043	AGRICOLA	19.400	-99.068	-0.205	-0.328	-0.301	-0.240	0.621	0.362	-0.215	-0.364	-0.156	-0.241	-0.320	-1.044	-0.977	-0.978				
704	6	IZTACALCI 06-044	AGRICOLA	19.395	-99.069	-0.901	-0.190	-0.094	0.323	1.008	0.742	-0.282	-0.095	-0.349	-0.241	-0.320	-1.055	-0.969	-0.969				
705	6	IZTACALCI 06-045	AGRICOLA	19.395	-99.062	-0.205	-0.328	-0.301	-0.240	1.421	1.260	-0.014	-0.311	0.008	-0.241	-0.320	-1.026	-1.000	-1.002				
706	6	IZTACALCI 06-046	AGRICOLA	19.388	-99.068	-0.901	-0.190	-0.094	0.323	0.703	0.357	0.655	-0.049	-0.549	-0.241	-0.320	-1.044	-0.998	-1.000				
707	6	IZTACALCI 06-047	AGRICOLA	19.388	-99.063	-0.205	-0.328	-0.301	-0.240	1.052	0.831	-0.482	-0.334	-0.084	-0.241	-0.320	-1.044	-0.976	-0.970				
708	6	IZTACALCI 06-002	AMPLIACI	19.387	-99.097	0.397	0.709	0.489	-0.211	0.632	0.609	0.320	0.412	-0.238	-0.241	-0.320	-0.884	-1.015	-1.017				
709	6	IZTACALCI 06-048	CAMPAMI	19.385	-99.114	-0.005	0.704	0.816	0.732	1.175	1.092	0.721	0.683	-0.100	-0.241	-0.320	-0.991	-1.004	-1.003				

4.4.2. Construcción del modelo multicriterio

El desarrollo de un modelo matemático con un enfoque multicriterio para la toma de decisiones enfrenta desafíos significativos. Uno de los retos es la selección de las variables que representen con precisión la accesibilidad a los servicios, en particular, a las escuelas de nivel primaria. El desafío más complejo radica en la asignación de pesos a estas variables, ya que la determinación de los pesos requiere un equilibrio adecuado que refleje su importancia relativa dentro del conjunto de datos en comparación con las demás variables.

Asimismo, la elección de una función de agregación adecuada es fundamental para integrar la información con los valores pesos de manera coherente y generar un modelo representativo de la accesibilidad. Por lo tanto, gestionar eficazmente las interacciones entre los distintos variables y sus efectos simultáneos resulta clave para el desarrollo de un modelo sólido y confiable. A continuación, se describen las actividades realizadas para la construcción del modelo matemático multicriterio.

1) Identificación de variables relevantes

La identificación de variables relevantes tiene como objetivo realizar una evaluación exhaustiva para seleccionar aquellas variables que influyen de manera significativa en el propósito del estudio (Gironés et al., 2017). Este proceso implica la aplicación de diversas técnicas analíticas que permitan determinar las variables de mayor impacto con el propósito de seleccionar un conjunto de variables que tengan un impacto significativo en el problema de la accesibilidad a servicios de las escuelas de nivel primaria.

Durante este proceso, se utilizaron de manera conjunta diversas técnicas estadísticas; ya que cada una proporciona una perspectiva única sobre los datos, lo que contribuye a una selección más completa y precisa de las variables. La combinación de estas técnicas permite optimizar la representatividad del modelo final, asegurando que refleje adecuadamente la realidad del problema analizado.

La aplicación conjunta de estas técnicas permitió abordar de manera integral los diversos aspectos del problema de accesibilidad; lo que facilitó una evaluación más completa y la obtención de resultados precisos y relevantes. En este contexto, la selección de variables se fundamentó en el uso complementario de distintas técnicas, asegurando un análisis riguroso y respaldado por enfoques metodológicos sólidos.

-
- o *Eliminación de redundancia*: La eliminación de variables que aportaban información similar. Esto permite simplificar el conjunto de datos y reducir la complejidad sin perder precisión en los resultados.
 - o *Análisis desde diversas perspectivas*: Identificación de las variables de mayor impacto desde diferentes enfoques. Esto con el propósito de obtener una comprensión más completa de la relación de las variables y alternativas en el contexto del problema.
 - o *Selección de variables*: Reducción del número de variables en aquellos que tienen un impacto significativo; asegurando que el análisis se centre en las variables más relevantes para la evaluación de la accesibilidad a servicios.

A continuación, se describen el proceso y los resultados de las técnicas implementadas.

Factor de Inflación de la Varianza

En esta investigación, el análisis mediante el Factor de Inflación de la Varianza permitió evaluar la multicolinealidad entre las variables con el fin de identificar y eliminar la redundancia de aquellos que aportan información similar, mediante la obtención del VIF. Aunque este enfoque es ampliamente utilizado en modelos de regresión, su aplicación resulta igualmente útil en la identificación de los variables relevantes, ya que permite detectar y gestionar la colinealidad entre variables, mejorando así la calidad y fiabilidad del modelo de análisis.

El proceso de cálculo consistió en determinar el valor del Factor de Inflación de la Varianza para cada variable, evaluando su relación con el resto de las variables dentro del conjunto de datos. Para llevar a cabo este análisis, se utilizó el lenguaje de programación R, específicamente la librería *car*, la cual proporciona herramientas especializadas para la detección de multicolinealidad y el análisis de regresión, garantizando así un enfoque preciso y eficiente en la evaluación de las variables.

El proceso de selección de variables relevantes se enfocó en eliminar aquellos que presentan multicolinealidad significativa y aportan información redundante. Se definió un umbral de Factor de Inflación de la Varianza superior a 10 como criterio para descartar variables altamente colineales. Además, se evaluaron variables con un Factor de Inflación de la Varianza entre 1 y 10; los cuales, aunque indicaban cierto grado de multicolinealidad, fueron retenidos debido a su importancia para el análisis del problema en estudio; ya que proporcionan información valiosa que no puede ser descartada sin afectar la interpretación y calidad del modelo.

Este análisis se llevó a cabo de manera detallada para cada una de las variables, evaluando tanto el valor del Factor de Inflación de la Varianza como la relevancia de cada criterio dentro del conjunto de datos, considerando su contribución a la explicación en el fenómeno de accesibilidad. La

Tabla 4.26 4.26 muestra una matriz que muestra los resultados del Factor de Inflación de la Varianza de cada criterio con el respecto del resto. Las celdas sombreadas indican los variables que presentan una alta colinealidad y, además, aportan información

redundante dentro del conjunto de datos. Muestra los variables seleccionados a partir del Factor de Inflación de la Varianza.

Tabla 4.26. Matriz del Factor de Inflación de la Varianza de cada criterio

	ANALF	SBASC	OVHAC	PSDSS	POBTOT	P6A11	P6A11NOA	PGRAPROES	PPEA	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	TIEM_PR_PIE	DIST_PR_PIE
ANALF		9.4369	7.8112	1.7189	16.3785	16.6892	1.0525	2.2796	1.2687	1.7573	1.7809	166.8374	16121.6589	14864.6914
SBASC	1.0515		2.7565	1.6423	16.3393	16.6445	1.0558	2.1395	1.2165	1.7593	1.8157	166.9285	16148.7127	14886.5058
OVHAC	1.0529	3.3347		1.7165	16.3662	16.7087	1.0570	2.2654	1.2417	1.7599	1.8121	166.9755	16157.1006	14895.2337
PSDSS	1.0522	9.0223	7.7950		16.4467	16.7980	1.0561	2.2796	1.2750	1.7574	1.8147	166.9309	16133.8462	14874.2780
POBTOT	1.0483	9.3854	7.7710	1.7197		1.6442	1.0565	2.2307	1.2367	1.7595	1.8138	162.1043	16141.3029	14891.0074
P6A11	1.0443	9.3476	7.7567	1.7172	1.6075		1.0580	2.2342	1.2374	1.7589	1.8136	166.2414	16133.6532	14878.2763
P6A11NOA	1.0475	9.4299	7.8035	1.7170	16.4281	16.8255		2.2796	1.2707	1.7599	1.8158	166.9818	16156.7673	14895.1148
PGRAPROES	1.0529	8.8687	7.7624	1.7201	16.0977	16.4908	1.0580		1.2683	1.7595	1.8164	166.8391	16157.1298	14895.4445
PPEA	1.0477	9.0164	7.6074	1.7201	15.9567	16.3298	1.0545	2.2677		1.7593	1.8160	166.8888	16157.1733	14895.3451
PVPH_SNBIE	1.0513	9.4461	7.8110	1.7176	16.4466	16.8157	1.0580	2.2790	1.2745		1.0814	166.8928	16155.7873	14894.6764
PVPH_SINTIC	1.0323	9.4454	7.7924	1.7184	16.4268	16.7995	1.0576	2.2796	1.2747	1.0477		166.9584	16157.0971	14895.4441
NUM_ESC	1.0520	9.4465	7.8110	1.7196	15.9705	16.7516	1.0580	2.2777	1.2743	1.7590	1.8162		13112.1652	13086.7310
TIEM_PR_PIE	1.0506	9.4447	7.8113	1.7176	16.4350	16.8018	1.0580	2.2796	1.2750	1.7598	1.8165	135.5130		130.0996
DIST_PR_PIE	1.0507	9.4439	7.8112	1.7177	16.4463	16.8069	1.0580	2.2796	1.2750	1.7599	1.8165	146.7065	141.1196	

Tabla 4.27. Variables seleccionadas a partir del factor de inflación de la varianza

ANALF	SBASC	OVHAC	PSDSS	P6A11	P6A11NOA
PGRAPROES	PPEA	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE

Correlación lineal

Otra de las técnicas que se empleó, es la técnica de correlación; la cual permite conocer el nivel de relación para poder seleccionar los que tiene una mayor significancia para incluir en el modelo de análisis. una correlación alta, para eliminar aquellos aportan la misma información. Se aplicó el coeficiente de correlación lineal al conjunto de datos. Dado que las variables no siguen una distribución normal, se utilizó el coeficiente de correlación de Spearman. Los utilizó el lenguaje de programación R, empleando las librerías corr, corrplot y PerformanceAnalytics para los cálculos, la visualización gráfica y análisis de los datos. Se estableció un umbral de correlación de ± 0.6 con el propósito de identificar relaciones significativas entre las variables. La Tabla 4.28 muestra la matriz de las correlaciones. Para una mayor comprensión, se muestran la matriz de diagramas de dispersión con los valores de las correlaciones y el mapa de calor de la matriz de correlación, véase Tabla 4.28 y Figura 4.33 respectivamente.

Tabla 4.28. Matriz de correlación

	ANALF	SBASC	OVHAC	PSDSS	POBTOT	P6A11	P6A11NOA	PGRAPROES	PPEA	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	TIEM_PR_PIE	DIST_PR_PIE
ANALF	1.0000	0.0081	0.0155	-0.0167	0.0316	-0.0002	0.1265	-0.0107	0.0694	0.0408	0.0815	-0.0707	-0.0664	-0.0620
SBASC	0.0081	1.0000	0.9253	0.6449	0.1012	0.3196	0.1021	0.8309	-0.4072	0.1963	0.2806	0.1501	0.1220	0.1186
OVHAC	0.0155	0.9253	1.0000	0.6024	0.1310	0.3418	0.1073	0.7885	-0.3377	0.2252	0.3210	0.1001	0.0692	0.0661
PSDSS	-0.0167	0.6449	0.6024	1.0000	0.1541	0.3001	0.1337	0.5403	-0.2741	0.1962	0.2606	0.0408	0.0497	0.0512
POBTOT	0.0316	0.1012	0.1310	0.1541	1.0000	0.9526	0.2286	0.0867	-0.0219	0.3053	0.3560	-0.5707	-0.3879	-0.3858
P6A11	-0.0002	0.3196	0.3418	0.3001	0.9526	1.0000	0.2472	0.3047	-0.1338	0.3386	0.4164	-0.4700	-0.3139	-0.3135
P6A11NOA	0.1265	0.1021	0.1073	0.1337	0.2286	0.2472	1.0000	0.0977	0.0257	0.0930	0.1587	-0.1135	-0.0575	-0.0557
PGRAPROES	-0.0107	0.8309	0.7885	0.5403	0.0867	0.3047	0.0977	1.0000	-0.4097	0.1988	0.2861	0.1168	0.0911	0.0881
PPEA	0.0694	-0.4072	-0.3377	-0.2741	-0.0219	-0.1338	0.0257	-0.4097	1.0000	-0.0999	-0.1254	-0.0627	-0.0463	-0.0452
PVPH_SNBIE	0.0408	0.1963	0.2252	0.1962	0.3053	0.3386	0.0930	0.1988	-0.0999	1.0000	0.6937	-0.1576	-0.0782	-0.0789
PVPH_SINTIC	0.0815	0.2806	0.3210	0.2606	0.3560	0.4164	0.1587	0.2861	-0.1254	0.6937	1.0000	-0.1748	-0.1030	-0.1038
NUM_ESC	-0.0707	0.1501	0.1001	0.0408	-0.5707	-0.4700	-0.1135	0.1168	-0.0627	-0.1576	-0.1748	1.0000	0.8198	0.8175
TIEM_PR_PIE	-0.0664	0.1220	0.0692	0.0497	-0.3879	-0.3139	-0.0575	0.0911	-0.0463	-0.0782	-0.1030	0.8198	1.0000	0.9986
DIST_PR_PIE	-0.0620	0.1186	0.0661	0.0512	-0.3858	-0.3135	-0.0557	0.0881	-0.0452	-0.0789	-0.1038	0.8175	0.9986	1.0000

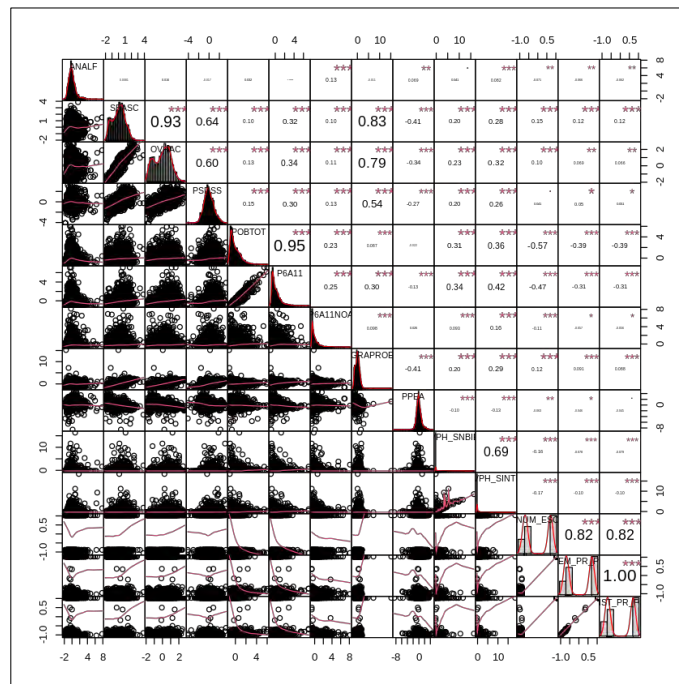


Figura 4.32. Matriz de diagramas de dispersión de la correlación

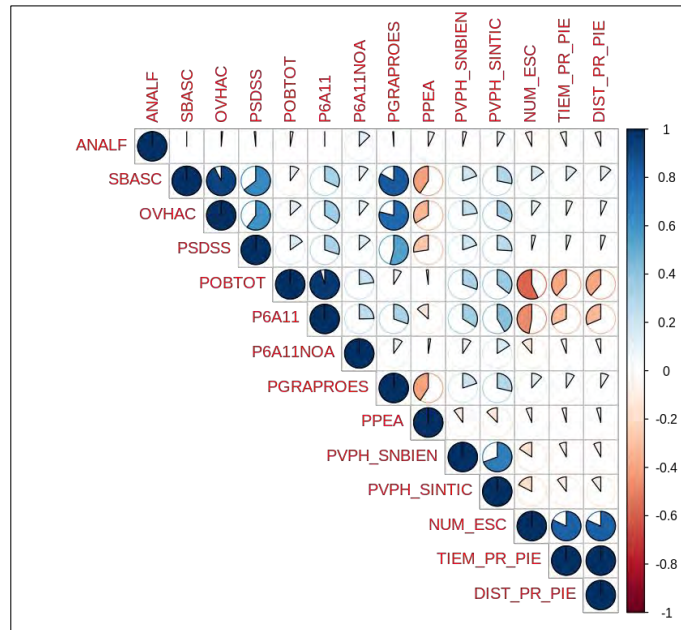


Figura 4.33. Mapa de calor de la correlación

Posteriormente, se aplicó una prueba de significancia t-Student a los valores que cumplen con el umbral establecido para saber si las variables son estadísticamente significativas. El valor de p obtenido es menor al nivel de significancia con $\alpha = 0.05$, rechazando la prueba de hipótesis y concluyendo que hay evidencia suficiente para afirmar que existe una correlación significativa fuerte entre ciertas variables. Este proceso se realizó para cada par de variables que mostraron una fuerte relación lineal. La ecuación 4.22 muestra el planteamiento de la prueba de hipótesis para la prueba de significancia. La Tabla 4.29 y Tabla 4.30 muestra los resultados de la prueba de significancia. Con base a los resultados.

La selección de variables se basó en aquellos valores que se encontraban dentro de dicho umbral. Además, se llevó a cabo un análisis exhaustivo de la contribución de cada criterio en el conjunto de datos, con el objetivo de identificar y eliminar aquellos que aportaban información redundante, a pesar de presentar una alta correlación. presenta los resultados de las correlaciones lineales obtenidas. La Tabla 4.31 muestra los variables seleccionados.

Planteamiento de la prueba de hipótesis para la prueba de significancia de la correlación lineal.

$$\begin{aligned}
 H_0: \rho &= 0 && \text{No hay una correlación significativa} \\
 &&& \text{entre las variables} \\
 H_1: \rho &\neq 0 && \text{Existe una correlación significativa} \\
 &&& \text{entre las variables}
 \end{aligned}
 \tag{4.22}$$

Tabla 4.29. Prueba de significancia de algunas variables. Parte 1

Spearman's rank correlation rho data: Data\$SBASC and Data\$OVHAC S = 73933494, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.9253143	Spearman's rank correlation rho data: Data\$SBASC and Data\$PSDSS S = 351488132, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.644936
Spearman's rank correlation rho data: Data\$PSDSS and Data\$OVHAC S = 393637398, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.6023579	Spearman's rank correlation rho data: Data\$POBTOT and Data\$P6A11 S = 46892491, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.9526304
Spearman's rank correlation rho data: Data\$SBASC and Data\$PGRAPROES S = 167402510, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.8308944	Spearman's rank correlation rho data: Data\$OVHAC and Data\$PGRAPROES S = 209332293, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.788538

Tabla 4.30. Prueba de significancia de algunas variables. Parte 2

Spearman's rank correlation rho data: Data\$PVPH_SNBIEN and Data\$PVPH_SINTIC S = 303191982, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.6937235	Spearman's rank correlation rho data: Data\$NUM_ESC and Data\$DIST_PR_PIE S = 180637706, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.8175246
Spearman's rank correlation rho data: Data\$NUM_ESC and Data\$TIEM_PR_PIE S = 178429036, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.8197557	Spearman's rank correlation rho data: Data\$DIST_PR_PIE and Data\$TIEM_PR_PIE S = 1367888, p-value < 2.2e-16 alternative hypothesis: true rho is not equal to 0 sample estimates: rho 0.9986182

Tabla 4.31. Variables seleccionadas a partir de la correlación lineal

SBASC	OVHAC	PSDSS	P6A11	PGRAPROES
PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE	

Análisis de componentes principales

Para la selección de las variables relevantes en este análisis, se consideró la información proporcionada por los valores propios y los vectores propios. Esta información permite evaluar la influencia de cada variable en cada vector o componente. En otras palabras, al considerar la proporción de la varianza explicada, determinada por los valores propios, la matriz de vectores propios (o matriz de

componentes principales) proporciona una comprensión más profunda de las relaciones subyacentes en los datos y facilita la identificación de las variables más influyentes en la estructura del conjunto de datos.

El primer paso consistió en aplicar el Análisis de Componentes Principales al conjunto de datos, determinando el porcentaje de varianza explicada como criterio para seleccionar las componentes más significativas. Se estableció un umbral inicial del 70% de varianza explicada. Como resultado se consideraron las primeras cinco componentes, como se muestra en la Figura 4.34. Los cálculos fueron realizados utilizando el lenguaje R, con las librerías *FactoMineR* y *factoextra*.

Posteriormente, se obtuvo la matriz de vectores propios y se seleccionaron los primeros cinco vectores según el porcentaje de varianza explicada. Considerando las propiedades de esta matriz, se procedió a normalizar los valores con el fin de representar la importancia de cada criterio en términos porcentuales. Se estableció un umbral de relevancia considerando las dos variables con los valores más altos por componente. La Tabla 4.32 presenta de manera porcentual la importancia relativa de cada criterio, mientras que la Figura 4.35 ofrece una representación visual. La Tabla 4.33 presenta las variables más relevantes obtenidas mediante la técnica de Análisis de Componentes Principales.

Importance of components:				
	PC1	PC2	PC3	PC4
Standard deviation	1.9302	1.8898	1.2507	1.08295
Proportion of Variance	0.2661	0.2551	0.1117	0.08377
Cumulative Proportion	0.2661	0.5212	0.6330	0.71672

Figura 4.34. Porcentaje de varianza explicada

Tabla 4.32. Importancia relativa de los variables por componente

	Dim.1	Dim.2	Dim.3	Dim.4
ANALF	0.002	0.061	6.770	39.470
SBASC	24.154	0.244	1.597	0.265
OVHAC	22.997	0.029	1.134	0.367
PSDSS	15.352	0.024	0.399	0.363
P6ATI	5.279	14.546	0.999	0.005
P6ATINOA	0.661	2.205	0.080	40.843
PGRAPROES	18.199	0.366	1.256	0.004
PPEA	6.343	0.265	2.132	10.947
PVPH_SNBIEN	2.948	0.685	41.525	5.098
PVPH_SINTIC	3.562	0.624	42.898	1.430
NUM_ESC	0.240	40.806	0.595	0.602
DIST_PR_PIE	0.262	40.145	0.615	0.607

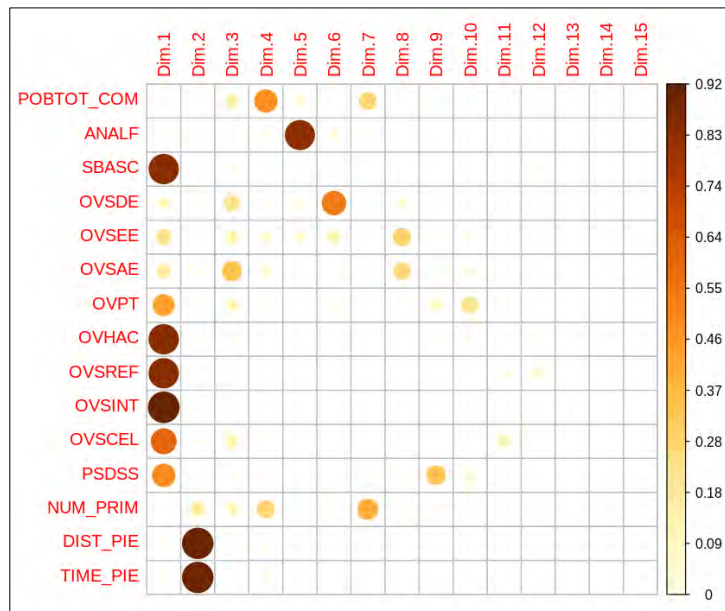


Figura 4.35. Gráfico de las Importancia relativa de las variables

Tabla 4.33. Variables seleccionadas de Análisis de Componentes Principales

ANALF	SBASC	OVHAC	P6ATI1NOA
PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE

Análisis factorial

Para determinar las variables relevantes del conjunto de datos, se aplicó el Análisis Factorial. En este punto, se tomó en cuenta la relación de los factores latentes con las variables; ya que estos factores representan estructuras subyacentes que explican la variabilidad de las variables.

Se realizó un análisis factorial al conjunto de datos con el objetivo de identificar las variables de mayor relevancia, a partir de los factores subyacentes (grupos de variables). En primer lugar, se aplicaron las pruebas estadísticas de Bartlett y Kaiser-Meyer-Olkin (KMO) para evaluar la idoneidad del análisis. La Ecuación 4.23 presenta el planteamiento de la prueba de hipótesis para el índice de adecuación muestral de KMO. El resultado obtenido fue un valor de KMO de 0.69, considerado como aceptable. Esto permite rechazar la hipótesis nula, indicando que existe suficiente varianza común entre las variables, lo que justifica la aplicación del análisis factorial. La Figura 4.36 muestra los resultados de la prueba KMO.

Posteriormente, se aplicó un análisis de sedimentación, junto con los criterios de Kaiser y el análisis paralelo de Horn, para determinar el número óptimo de factores. Estas técnicas permitieron una selección fundamentada del número de factores a retener en el análisis. La Figura 4.37 muestra el número de factores óptimo de factores. Para la obtención del análisis factorial se utilizó el método de rotación varimax, el cual busca simplificar la interpretación de los factores al maximizar la varianza de los cuadrados de las cargas factoriales. También, se empleó el método de estimación de

factores mediante mínimos cuadrados residuales. Se utilizó la correlación de spearman, más apropiada para variables no siguen una distribución normal; y se aplicó una matriz de covarianza robusta que resulta menos sensible a valores atípicos. Todos los cálculos se realizaron en el lenguaje de programación R, empleando las librerías nFactors, psych y MASS. La Figura 4.38 muestra los valores de las cargas factoriales; mientras que la Figura 4.39 presenta el grafo que ilustra la relación entre las variables observables y los factores subyacentes.

Como criterio se tomaron en cuenta las estimaciones de las cargas factoriales, las cuales indican la fuerza y dirección de la relación entre cada variable observada y los factores latentes. Además, se consideraron las variables de comunalidad y residuo. El criterio de comunalidad evalúa cuánta varianza de la variable es explicada por el factor correspondiente, mientras que el criterio de residuo específico indica cuánta varianza de la variable no es explicada por el factor. En la selección de variables, se eligieron las variables con las tres cargas factoriales más altas en cada factor, véase Tabla 4.34.

Planteamiento de la prueba de hipótesis para la prueba de Kaiser-Meyer-Olkin

H_0 : No hay suficiente varianza común entre las variables observadas. (4.23)

H_1 : Existe suficiente varianza común entre las variables observadas.

Kaiser-Meyer-Olkin factor adequacy					
Call: KMO(r = Data)					
Overall MSA = 0.69					
MSA for each item =					
ANALF	SBASC	OVHAC	PSDSS	P6A11	P6A11NOA
0.48	0.72	0.74	0.95	0.73	0.74
PGRAPROES	PPEA	PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
0.94	0.83	0.57	0.58	0.51	0.52

Figura 4.36. Resultados de la prueba de Kaiser-Meyer-Olkin

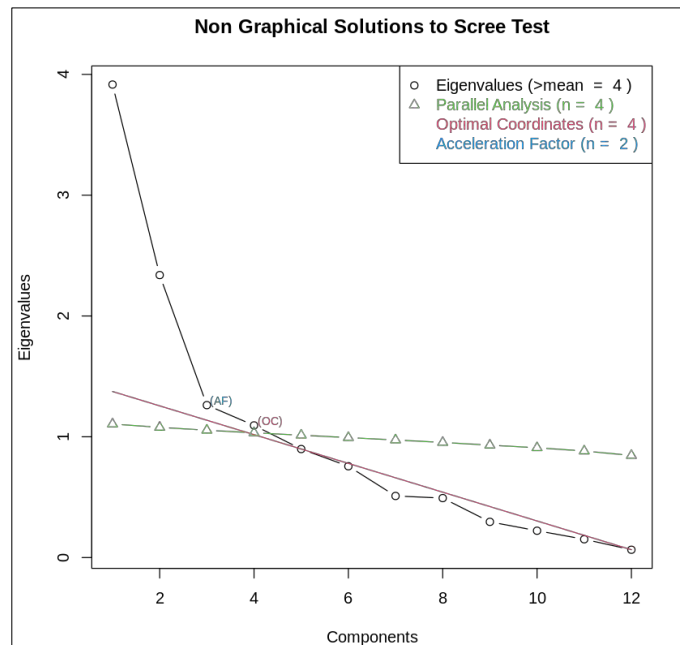


Figura 4.37. Número óptimo de factores

Standardized loadings (pattern matrix) based upon correlation matrix						
	MR1	MR2	MR3	h2	u2	com
ANALF	-0.02	-0.06	0.08	0.0095	0.990	2.0
SBASC	0.98	0.05	0.08	0.9780	0.022	1.0
OVHAC	0.91	0.01	0.14	0.8449	0.155	1.0
PSDSS	0.64	-0.01	0.15	0.4331	0.567	1.1
P6A11	0.33	-0.44	0.37	0.4421	0.558	2.8
P6A11NOA	0.11	-0.12	0.15	0.0504	0.950	2.7
PGRAPROES	0.85	0.03	0.12	0.7297	0.270	1.0
PPEA	-0.41	-0.03	-0.04	0.1726	0.827	1.0
PVPH_SNBIEI	0.14	-0.05	0.73	0.5511	0.449	1.1
PVPH_SINTIC	0.21	-0.06	0.90	0.8561	0.144	1.1
NUM_ESC	0.12	0.99	-0.16	1.0138	-0.014	1.1
DIST_PR_PIE	0.09	0.79	-0.07	0.6327	0.367	1.0
	MR1	MR2	MR3			
SS loadings	3.30	1.82	1.59			
Proportion Var	0.27	0.15	0.13			
Cumulative Var	0.27	0.43	0.56			
Proportion Explained	0.49	0.27	0.24			
Cumulative Proportion	0.49	0.76	1.00			

Figura 4.38. Resultados del Análisis Factorial

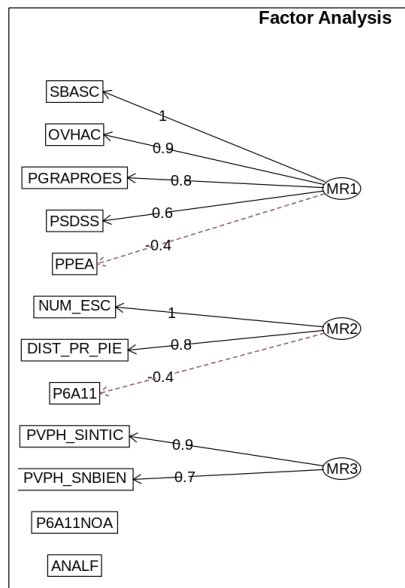


Figura 4.39. Resultados del Análisis Factorial de manera gráfica

Tabla 4.34. Variables seleccionadas a partir del Análisis Factorial

SBASC	OVHAC	P6A11	PGRAPROES
PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE

Selección de variables

A partir de las técnicas implementadas, se seleccionaron aquellas variables que demostraron una relevancia significativa en cada uno de los métodos utilizados, priorizando aquellas que mostraron mayor consistencia y relevancia a lo largo de todas las técnicas empleadas. En la Tabla 4.35 se presentan los resultados de las variables seleccionadas para la construcción del modelo multicriterio.

En consecuencia, la matriz de decisión; es decir el conjunto de datos para la formulación del modelo, se compone de 7 variables y 1,811 instancias o alternativas (localidades). Las variables seleccionadas fueron organizadas en tres dimensiones o criterios: educación, vivienda y espacial. Cada uno de estos grupos representa un aspecto clave en la evaluación de la accesibilidad a los servicios educativos, como se muestra en la Figura 4.40. Variables relevantes clasificados por dimensión.

Esta clasificación permitió un análisis estructurado que facilita la identificación de patrones de accesibilidad en función de distintas variables interrelacionadas.

Tabla 4.35. Variables relevantes de mayor impacto

Variables seleccionadas	
Factor de Inflación de la Varianza	ANALF, SBASC, OVHAC, PSDSS, P6A11, P6A11NOA, PGRAPROES, PPEA, PVPH_SNBIEN, PVPH_SINTIC, NUM_ESC, DIST_PR_PIE
Correlación lineal	SBASC, OVHAC, PSDSS, P6A11, PGRAPROES, PVPH_SNBIEN, PVPH_SINTIC, NUM_ESC, DIST_PR_PIE
Análisis de Componentes Principales	ANALF, SBASC, OVHAC, P6A11NOA, PVPH_SNBIEN, PVPH_SINTIC, NUM_ESC, DIST_PR_PIE
Análisis Factorial	SBASC, OVHAC, P6A11, PGRAPROES, PVPH_SNBIEN, PVPH_SINTIC, NUM_ESC, DIST_PR_PIE
Variables relevantes	SBASC, OVHAC, PGRAPROES, PVPH_SNBIEN, PVPH_SINTIC, NUM_ESC, DIST_PR_PIE

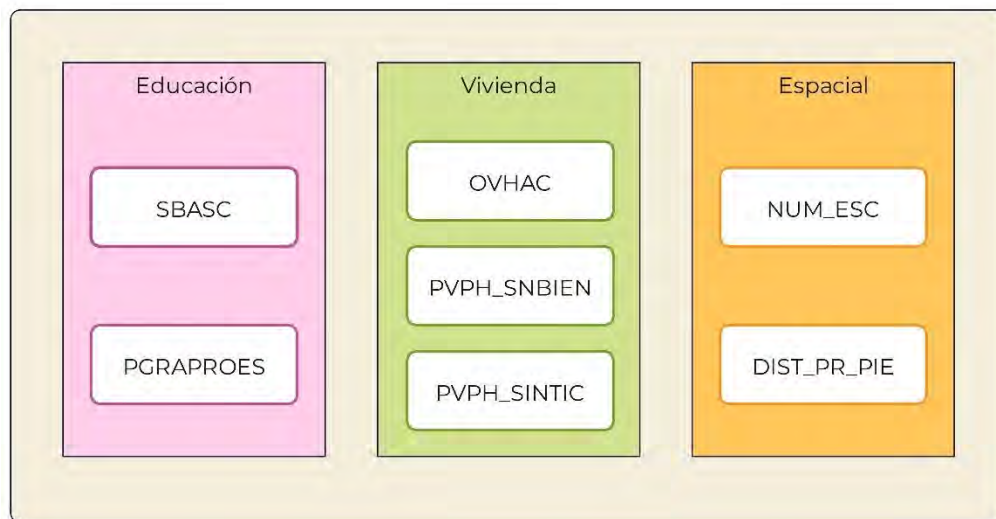


Figura 4.40. Variables relevantes clasificados por dimensión

2) Asignación de pesos de las variables y Formulación del modelo multicriterio

La asignación de pesos a las variables consiste en otorgar valores numéricos a cada una de las variables, de acuerdo con su importancia relativa dentro del conjunto de datos. En este contexto, los pesos reflejan el grado de relevancia o influencia de cada variable en relación con las demás. Para determinar estos valores, es necesario emplear un método de ponderación de variables que permita capturar adecuadamente el impacto de cada variable en la decisión final.

La asignación de pesos desempeña un papel fundamental en el análisis de decisiones multicriterio; ya que facilita la comparación de alternativas que presentan conflictos con el objetivo de seleccionar la mejor opción. Este proceso implica obtener los valores de los pesos de cada variable, para poder establecer una jerarquización o ranking de las variables. Sin embargo, uno de los principales retos de este procedimiento radica en determinar los valores adecuados que representen, de manera precisa, el nivel de importancia de cada criterio dentro del análisis (Malczewski, 2017), (Odu, 2019).

En esta investigación se utilizó el método de ponderación PCA-Entropía con enfoque objetivo propuesto; el cual combina técnicas matemáticas avanzadas para asignar los pesos de las variables. La Figura 4.41 muestra un esquema general del procedimiento seguido en la aplicación del método PCA-Entropía desarrollado en la sección 4.2.

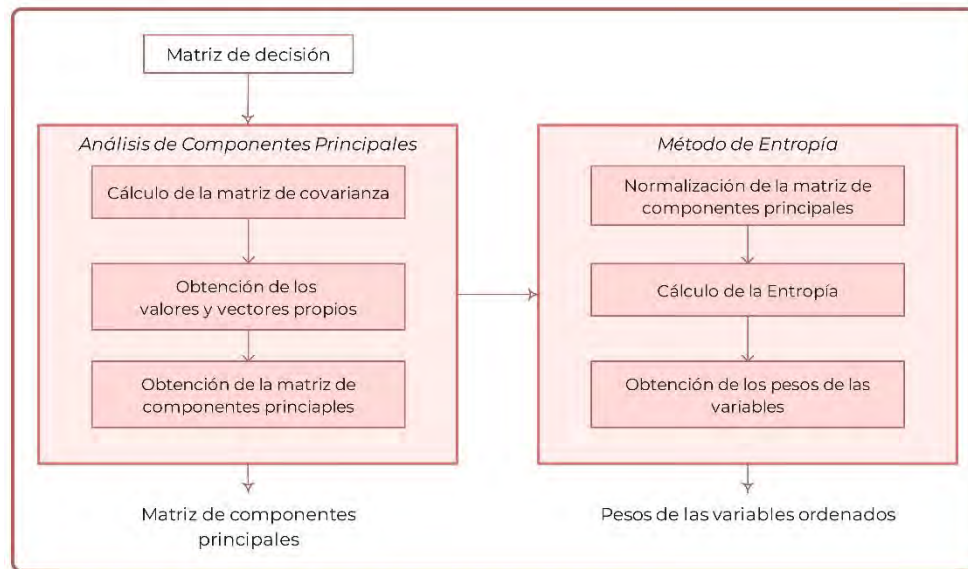


Figura 4.41. Método ponderación de variables PCA-Entropía

Por otro lado, la formulación del modelo multicriterio se enfoca en definir una función matemática que integre de manera sistemática, estructurada y equilibrada los pesos asignados a las variables junto con sus respectivos valores de las variables correspondientes en cada alternativa. Una vez definidos los pesos, el siguiente paso consiste en establecer la función de agregación para obtener el modelo matemático multicriterio. Cabe señalar que el modelo de agregación puede adoptar distintas estructuras, dependiendo del tipo de función de agregación seleccionada, con el objetivo de reflejar de manera adecuada la contribución relativa de cada criterio en el proceso de evaluación.

Con el objetivo de obtener resultados más robustos y precisos en la fase de construcción del modelo multicriterio, se incorporó la técnica de validación cruzada integrando de manera conjunta las etapas de asignación de pesos a las variables y la formulación del modelo. La validación cruzada es ampliamente reconocida como un procedimiento eficaz para evaluar el rendimiento y la capacidad de generalización de un modelo.

La validación cruzada consiste en dividir el conjunto de datos en múltiples subconjuntos o particiones; donde en cada iteración el modelo se entrena y evalúa con una parte de los datos; al final se evalúa con la parte restante. Al repetir este proceso varias veces con diferentes particiones del conjunto de datos, se obtiene una estimación más precisa y confiable del desempeño del modelo bajo distintas condiciones (Hernández, et al, 2004).

En esta investigación se empleó la técnica de validación cruzada k-fold, debido a su efectividad para equilibrar el sesgo y la varianza en la evaluación de modelos. Este enfoque consiste en dividir el conjunto de datos en k subconjuntos (o pliegues) de tamaño similar, realizando posteriormente k iteraciones. En cada iteración, uno de los subconjuntos (pliegues) se utiliza como conjunto de prueba, mientras que los pliegues restantes se emplean para entrenar el modelo.

La integración de la validación cruzada con las etapas de asignación de pesos a las variables y la formulación del modelo multicriterio resulta particularmente relevante, ya que contribuye a mejorar la precisión del modelo al identificar la combinación óptima de variables ponderadas, minimizando tanto el sesgo como la varianza en las estimaciones.

Otro beneficio es que la evaluación repetida sobre distintos subconjuntos de datos fortalece la capacidad de generalización del modelo; asegurando su aplicabilidad en nuevos contextos y reduciendo el riesgo de sobreajuste. Este procedimiento disminuye la dependencia de una única partición de datos; proporcionando una estimación más robusta, estable, confiable y representativa del rendimiento real del modelo. Asimismo, un aspecto importante de la validación cruzada es el incremento en la confiabilidad del modelo multicriterio, al ofrecer una evaluación rigurosa y sistemática de la metodología desarrollada, lo cual garantiza una mayor consistencia y solidez en los resultados obtenidos.

La validación cruzada k-fold, integrada con las fases de ponderación de variables y la formulación del modelo multicriterio, constituyó un proceso estructurado compuesto por diversas etapas, las cuales se detallan a continuación.

2.1) Partición del conjunto de datos

Una vez transformado el conjunto de datos e identificadas las variables relevantes correspondientes a los criterios de educación, vivienda y dimensión espacial; el primer paso consistió en la partición del conjunto de datos en dos subconjuntos: uno para el entrenamiento del modelo y otro para su validación. De un total de 1,811 instancias disponibles, se asignó el 75 % (1,357 instancias) al conjunto de entrenamiento y el 25 % restante (454 instancias) al conjunto de validación.

La proporción de esta partición fue determinada con base en diversos experimentos preliminares, en los que se evaluaron distintas configuraciones para asegurar un equilibrio adecuado entre la capacidad de ajuste del modelo y su potencial de generalización.

2.2) Configuración de la Validación Cruzada k-Fold

En este proceso se utilizó el conjunto de entrenamiento conformado por 1,357 instancias, obtenido en la etapa anterior. Sobre este conjunto se aplicó la técnica de validación cruzada k-fold con 5 pliegues; lo que implicó dividir los datos en cinco subconjuntos de tamaño similar. En cada iteración, se utilizaron cuatro pliegues (1,088 instancias) para entrenar el modelo y uno (269 instancias) para evaluar su rendimiento.

Se realizaron múltiples experimentos variando tanto la partición de los datos como el número de pliegues, determinándose que la configuración con 5 pliegues ofrecía el mejor equilibrio entre estabilidad y capacidad de generalización del modelo.

Los cálculos y análisis se llevaron a cabo utilizando el lenguaje de programación R, con el apoyo de las librerías dplyr, FactoMineR, factoextra, classInt, writexl y caTools.

2.2.1) Asignación de pesos de las variables

En esta etapa se implementó el método ponderación PCA-Entropía propuesto, con el propósito de obtener los coeficientes que cuantifican la importancia relativa de cada variable.

Paso 1. Cálculo de la matriz de covarianza.

Como primer paso se calculó la matriz de covarianza para identificar el grado de relación lineal existente entre cada par de variables del conjunto de datos. La Tabla 4.36 presenta la matriz de covarianza resultante de este análisis.

Tabla 4.36. Matriz de covarianza de la iteración 1

	ANALF	SBASC	OVSDE	OVSEE	OVSAE	OVPT	OVHAC	OVSREF
ANALF	0.965	0.020	-0.043	0.099	-0.047	0.017	0.031	-0.020
SBASC	0.020	1.022	0.201	0.331	0.278	0.534	0.948	0.830
OVSDE	-0.043	0.201	1.092	0.206	0.312	0.261	0.235	0.279
OVSEE	0.099	0.331	0.206	1.140	0.365	0.498	0.351	0.461
OVSAE	-0.047	0.278	0.312	0.365	1.141	0.506	0.278	0.487
OVPT	0.017	0.534	0.261	0.498	0.506	1.098	0.536	0.653
OVHAC	0.031	0.948	0.235	0.351	0.278	0.536	1.016	0.849
OVSREF	-0.020	0.830	0.279	0.461	0.487	0.653	0.849	1.009

Paso 2. Cálculo de los valores y vectores propios

En este paso se calcularon los valores propios y los vectores propios. Los valores propios explican el porcentaje de varianza explicada. Los vectores propios reflejan las direcciones principales del espacio de los datos, representando las combinaciones lineales de las variables originales. La Tabla 4.37 muestra los valores propios obtenidos en la iteración 1. La Tabla 4.38 muestran la matriz de vectores propios obtenida en la iteración 1, también conocida como matriz de componentes principales.

Tabla 4.37. Vector de valores propios de la iteración 1

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
3.907	1.212	1.039	0.875	0.720	0.504	0.158	0.068
0.4606	0.6035	0.7259	0.8290	0.9140	0.9733	0.99197	1

Tabla 4.38. Matriz de vectores propios de la iteración 1

CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8
0.009	0.137	0.798	0.466	0.353	-0.004	0.042	-0.013
0.434	0.405	-0.120	0.084	0.020	0.133	-0.424	-0.654
0.208	-0.488	-0.282	0.772	-0.199	-0.052	-0.009	-0.024
0.315	-0.278	0.487	-0.240	-0.673	0.267	-0.069	-0.005
0.309	-0.552	-0.019	-0.258	0.612	0.365	-0.156	0.027
0.405	-0.148	0.110	-0.204	0.069	-0.864	-0.095	0.019
0.439	0.387	-0.112	0.115	-0.003	0.140	-0.234	0.746
0.466	0.156	-0.084	-0.031	0.055	0.097	0.852	-0.114

Paso 3. Obtención de la matriz de componentes principales

Para asegurar una representación adecuada en función de la variabilidad en los datos, se estableció un umbral del 90% de varianza explicada; con base a la información de los valores propios. Por tal motivo se seleccionaron las primeras componentes de la matriz de vectores propio; es decir, los primeros cinco vectores propios. El resultado es una matriz de decisión de dimensión 8×5. La Tabla 4.39 muestra la matriz de componentes principales seleccionada en la iteración 1.

Tabla 4.39. Matriz de componentes principales de la iteración 1

CP1	CP2	CP3	CP4	CP5
0.009	0.137	0.798	0.466	0.353
0.434	0.405	-0.120	0.084	0.020
0.208	-0.488	-0.282	0.772	-0.199
0.315	-0.278	0.487	-0.240	-0.673
0.309	-0.552	-0.019	-0.258	0.612
0.405	-0.148	0.110	-0.204	0.069
0.439	0.387	-0.112	0.115	-0.003
0.466	0.156	-0.084	-0.031	0.055

Paso 4. Normalización de la matriz de componente principales seleccionada

Considerando las propiedades matemáticas de la matriz de vectores propios seleccionada; como la independencia lineal, la generación del espacio vectorial, la ortogonalidad y la normalización (vectores ortonormales), el proceso de normalización aplicado consistió en elevar al cuadrado cada elemento, obteniendo así vectores normalizados cuya suma por columna es igual a 1. La Tabla 4.40 presenta la matriz de componentes principales seleccionada y normalizada.

Tabla 4.40. Matriz de componentes principales normalizada de la iteración 1

CP1	CP2	CP3	CP4	CP5
0.00009	0.01867	0.63682	0.21757	0.12489
0.18816	0.16409	0.01438	0.00706	0.00040
0.04344	0.23841	0.07954	0.59560	0.03960
0.09903	0.07754	0.23706	0.05738	0.45270
0.09521	0.30516	0.00035	0.06658	0.37458
0.16424	0.02185	0.01210	0.04156	0.00481
0.19274	0.14988	0.01261	0.01326	0.00001
0.21709	0.02441	0.00714	0.00098	0.00301

Paso 5. Cálculo de la Entropía

En el siguiente paso se calculó la entropía de cada componente de la matriz. La Tabla 4.41 muestra el vector de los valores de entropía para cada componente.

Tabla 4.41. Cálculo de la Entropía de cada componente de la iteración 1

CP1	CP2	CP3	CP4	CP5
0.188	0.191	0.210	0.205	0.206

Paso 6. Obtención de los pesos de las variables

Empleando los cálculos correspondientes, en esta etapa se obtuvieron los pesos de los valores de los pesos de las variables. La Tabla 4.42 muestra los pesos obtenidos para cada una de las variables mostradas en un ranking en la iteración 1.

Tabla 4.42. Pesos de los variables obtenidos en la iteración 1

	SBASC	OVHAC	PGRAPROES	PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
Pesos	0.0853	0.0860	0.0660	0.2600	0.2549	0.1241	0.1237
Ranking	6	5	7	1	2	3	4

Los pesos obtenidos, mediante el método de ponderación de variables PCA-Entropía, representan los coeficientes que serán utilizados para la formulación del modelo multicriterio orientado a evaluar la accesibilidad a los servicios.

2.2.2) Formulación del modelo

Una vez determinados los pesos de las variables, se procedió a la formulación del modelo matemático multicriterio.

Paso 1. Definición del modelo matemático de agregación

En este punto del estudio se determinó que el modelo más adecuado es la función lineal de suma ponderada, con base en diversos experimentos y pruebas realizadas. Este modelo destaca por su facilidad de interpretación y capacidad para representar de forma clara y coherente la relación entre las variables y los objetivos del análisis. La Ecuación 4.24 muestra la formulación matemática correspondiente

al modelo de agregación lineal mediante suma ponderada, aplicada en la iteración 1.

$$f(\bar{x}) = 0.0853x_1 + 0.0860x_2 + 0.0660x_3 + 0.2600x_4 + 0.2549x_5 + 0.1241x_6 + 0.1237x_7 \quad (4.24)$$

Donde:

IAS Modelo de accesibilidad a servicios

x_1 SBASC, Población de 15 años o más sin educación básica

x_2 OVHAC, Viviendas con hacinamiento

x_3 PGRAPROES, Grado promedio de escolaridad

x_4 PVPH_SNBIEN, Viviendas particulares habitadas sin ningún bien

x_5 PVPH_SINTIC, Viviendas particulares habitadas sin tecnologías de la información y de la comunicación (TIC)

x_6 NUM_ESC, Número de escuelas ubicadas en la localidad

x_7 DIST_PR_PIE, Distancia promedio del origen al destino, en la modalidad de transporte a pie

Paso 2. Evaluación y selección de las alternativas

En el punto de evaluación, cada alternativa fue evaluada utilizando el modelo matemático obtenido, aplicándolo a los datos de entrenamiento correspondientes de la iteración 1 (1,088 instancias). Como resultado de esta evaluación, se obtuvo un puntaje global para cada alternativa. Dicho puntaje permitió ordenar las alternativas en un ranking, de menor a mayor valor, con base en su desempeño según el modelo. Este ordenamiento facilitó la identificación del desempeño relativo de cada alternativa conforme a los criterios establecidos.

Para la selección de alternativas se utilizó el método de optimización de intervalos de Jenks, cuyo objetivo es minimizar la varianza interna dentro de cada grupo y maximizar las diferencias entre los distintos grupos. Este método resulta especialmente apropiado para la segmentación de datos, ya que garantiza que los grupos generados sean internamente homogéneos y, a la vez, claramente diferenciables entre sí. La aplicación de este enfoque permitió clasificar eficientemente los puntajes obtenidos por el modelo en cinco categorías, asegurando tanto la coherencia interna como una separación efectiva entre clases.

Dado que no existe una variable directa o estandarizada que mida específicamente la accesibilidad a los servicios, se recurrió al uso del Índice de Marginación como variable de referencia para la comparación. Este índice, elaborado por el Consejo Nacional de Población (CONAPO), constituye una medida consolidada que refleja las carencias sociales de una zona geográfica, al considerar dimensiones relacionadas con la educación, las condiciones de la vivienda, y el acceso a bienes y servicios. Asimismo, se presenta con un nivel de desagregación por localidad, lo que permite establecer comparaciones directas con los resultados obtenidos por el modelo propuesto. Su carácter oficial, reconocimiento institucional y enfoque multidimensional lo convierten en un criterio pertinente y confiable para evaluar el grado de concordancia del modelo con la realidad socioeconómica de las localidades; así como para identificar tendencias, patrones y similitudes espaciales entre el indicador propuesto y las condiciones sociales observadas.

El índice de marginación clasifica las localidades en cinco niveles: (1) Muy baja, (2) Baja, (3) Media, (4) Alta y (5) Muy alta marginación. Siguiendo esta misma lógica, el indicador de accesibilidad generado fue clasificado en cinco categorías, lo que permitió realizar una comparación más directa y estructurada entre ambas medidas. La Tabla 4.43 presenta ejemplos representativos de localidades correspondientes a la Iteración 1, incluyendo los valores de las variables, los resultados obtenidos mediante la evaluación del modelo y su respectiva clasificación en una de las cinco categorías.

Tabla 4.43. Subconjunto de datos evaluados y clasificados en la iteración 1

ID	SBASC	OVHAC	PGRAPROES	PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE	Datos evaluados	Datos clasificados
21	1.6725	0.3854	0.5913	-0.2411	-0.3201	-0.9908	-1.0273	-0.3979	5
237	0.0477	-0.7007	-0.3606	-0.2411	-0.3201	1.0360	1.0345	-0.0293	4
1530	-0.3815	1.4658	1.1782	-0.2411	-0.3201	-0.8841	-0.0727	0.0444	4
700	-0.9013	-0.1898	-0.1712	-0.2411	-0.3201	-0.9908	-0.8902	-0.4129	5
1233	-0.0990	0.1554	0.5642	-0.2411	-0.3201	1.0360	1.0345	0.1831	3
271	0.9545	-0.9661	-0.5693	-0.2411	-0.3201	1.0360	1.0345	-0.0907	4
1804	-0.8742	-1.2730	-1.0825	-0.2411	-0.3201	1.0360	1.0345	-0.1646	4
1584	-0.3060	0.6511	0.6926	-0.2411	-0.3201	-0.8841	-1.0080	-0.1871	4
513	-0.0609	0.7643	0.2176	-0.2411	-0.3201	1.0360	1.0345	0.2262	3
422	0.5361	0.6129	-0.0164	-0.2411	-0.3201	-0.8841	-1.0000	-0.2429	4
345	0.1982	-1.3210	-1.3006	-0.2411	-0.3201	-0.8841	-1.0510	-0.6760	5
660	-0.3231	0.7361	0.1715	-0.2411	-0.3201	1.0360	1.0345	0.2800	3
1432	0.3115	0.8359	2.0755	-0.2411	-0.3201	-0.8841	-1.0613	-0.1055	4
1356	-0.1950	0.0306	0.9241	-0.2411	-0.3201	1.0360	1.0345	0.2301	3

Paso 3. Evaluación del rendimiento de modelo

El siguiente paso consistió en evaluar el rendimiento del modelo propuesto utilizando el conjunto de prueba, conformado por 269 instancias dentro del proceso de validación cruzada. Este procedimiento implicó incorporar los valores de las variables de cada instancia en la ecuación del modelo multicriterio (Ecuación 4.1), con el fin de obtener su respectivo puntaje global. Posteriormente, se aplicó el método de optimización de clasificación por intervalos de Jenks para segmentar dichos puntajes en cinco categorías. La Tabla 4.44 presenta algunos ejemplos de localidades correspondientes al conjunto de prueba de la Iteración 1; en la que se muestran los valores de las variables, los puntajes generados por el modelo y la categoría asignada a cada localidad.

Dado que no existe una variable de referencia directa para evaluar el rendimiento del modelo de accesibilidad, se utilizó el coeficiente de correlación de Spearman como métrica principal. Este coeficiente permitió examinar la consistencia, coherencia y el grado de concordancia entre los valores generados por el modelo y los del índice de marginación. A partir de este análisis, fue posible identificar patrones comunes y tendencias similares entre los niveles de accesibilidad estimados por el modelo y las carencias sociales representadas en dicho índice. Además, se aplicó una prueba de significancia estadística con un nivel de confianza del 95% ($\alpha = 0.05$), con el fin de comprobar que la correlación observada no fuera producto del azar. Los resultados correspondientes a la Iteración 1 se presentan en la Tabla 4.45.

Este procedimiento se repite de manera análoga en cada uno de los pliegues generados durante el proceso de validación cruzada. Es decir, en cada iteración, un

subconjunto diferente de datos es utilizado como conjunto de prueba, mientras que los datos restantes se emplean para entrenar el modelo. Posteriormente, se evalúa el desempeño del modelo sobre el conjunto de prueba correspondiente, aplicando la misma métrica de evaluación previamente establecida.

Tabla 4.44. Subconjunto de datos de prueba evaluados y clasificados en la iteración 1

FID	SBASC	OVHAC	PGRAPROES	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE	Evaluación en IAS	Clase Jenks
1121	-0.91286	-1.77664	-0.96316	-0.24110	-0.32013	1.03604	1.03453	-0.18207	2
34	-1.54291	-1.25581	-1.31982	-0.24110	0.57825	1.03604	1.03453	0.01442	2
693	-0.73235	-1.10063	-0.38299	-0.24110	-0.32013	-0.99082	-1.03109	-0.57719	1
281	-0.47784	-0.17878	-1.44004	-0.24110	-0.32013	1.03604	1.03453	-0.03900	2
1244	0.23088	0.01656	-0.02253	-0.24110	-0.32013	-0.88414	-1.03613	-0.36249	2
792	1.25454	1.14187	0.86478	0.83251	0.23511	1.03604	1.03453	0.79522	4
741	0.43012	0.58716	0.10371	-0.24110	-0.32013	-0.99082	-0.99173	-0.29583	2
499	1.26256	1.31633	1.14920	-0.24110	-0.32013	-1.04415	-0.96182	-0.09598	2
1127	-1.97868	-1.70240	-0.98854	-0.24110	-0.32013	-0.99082	-0.92887	-0.76264	1
1558	-0.28489	-0.42571	-0.35210	0.54662	0.22305	-0.88414	-0.87702	-0.10338	2
726	0.33629	0.54477	-0.05349	-0.24110	-0.32013	1.03604	1.03453	0.18424	3
654	-0.57621	-1.25705	-0.35224	-0.24110	-0.32013	-0.88414	-1.01451	-0.56000	1
1208	0.97787	0.91907	0.98769	-0.24110	-0.32013	1.03604	1.03453	0.33991	3
1251	0.44604	-0.04320	0.04282	-0.24110	-0.32013	1.03604	1.03453	0.14939	3
1398	0.31716	0.56111	0.38850	-0.24110	-0.32013	1.03604	1.03453	0.21319	3
1232	-1.25563	-1.55896	-0.61406	-0.24110	-0.32013	1.03604	1.03453	-0.16956	2
1276	2.41471	1.25559	2.43449	-0.24110	-0.32013	1.03604	1.03453	0.58698	3

Tabla 4.45. Métrica de la evaluación del rendimiento de la iteración 1

Iteración	Coefficiente de correlación de Spearman	Prueba de significancia
1	-0.49030	$5.610571178134415 \times 10^{-18}$

Paso 4. Integración de la información

Una vez concluido el proceso de validación cruzada, se procedió a calcular el promedio de los pesos obtenidos en cada una de las cinco iteraciones. Este proceso permitió determinar los coeficientes finales que definen al modelo matemático multicriterio de accesibilidad a servicios. Dichos coeficientes representan el nivel de influencia relativa de cada variable dentro del modelo, garantizando mayor estabilidad y robustez en los resultados al considerar las variaciones observadas en cada iteración del proceso de validación. La Tabla 4.46 muestra los pesos asignados a cada una de las variables consideradas, obtenidos a lo largo de las cinco iteraciones realizadas y los valores del promedio de los pesos.

Por su parte, la Tabla 4.47 presenta los modelos multicriterio obtenidos en cada iteración, formulados mediante una función de agregación lineal con suma ponderada. En la Tabla 4.48 se muestran los resultados de la métrica de evaluación del rendimiento del modelo en cada iteración, definida a partir del coeficiente de correlación de Spearman; así como los resultados de las pruebas de significancia aplicadas. En general, los valores obtenidos para el coeficiente de correlación muestran una relación negativa de magnitud moderada. Esta información indica una tendencia inversa entre los valores del modelo multicriterio de accesibilidad a servicios y los del índice de marginación; es decir, a medida que la accesibilidad aumenta, los niveles de marginación tienden a disminuir, y viceversa. Este comportamiento observado refleja coherencia en las características

socioeconómicas de las localidades; desde el punto de vista de la relación entre accesibilidad y marginación.

Tabla 4.46. Pesos obtenidos en cada iteración y pesos promedio

Iteración	-	SBASC	OVHAC	PGRAPROES	PVPH_SNBIEN	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
1	Pesos	0.0853	0.0860	0.0660	0.2600	0.2549	0.1241	0.1237
	Ranking	6	5	7	1	2	3	4
2	Pesos	0.1146	0.1294	0.2502	0.0608	0.1962	0.1245	0.1241
	Ranking	6	3	1	7	2	4	5
3	Pesos	0.1159	0.1303	0.2529	0.0831	0.1675	0.1253	0.1249
	Ranking	6	3	1	7	2	4	5
4	Pesos	0.1143	0.1315	0.2533	0.1382	0.1148	0.1240	0.1238
	Ranking	7	3	1	2	6	4	5
5	Pesos	0.1144	0.1298	0.2488	0.1256	0.1293	0.1260	0.1261
	Ranking	7	2	1	6	3	5	4
Promedio	Pesos	0.1089	0.1214	0.2143	0.1335	0.1726	0.1248	0.1245
	Ranking	7	6	1	3	2	4	5

Tabla 4.47. Modelos multicriterio obtenidos en cada iteración

Iteración	Modelo matemático multicriterio
1	$f(\vec{x}) = 0.0853x_1 + 0.0860x_2 + 0.0660x_3 + 0.2600x_4 + 0.2549x_5 + 0.1241x_6 + 0.1237x_7$
2	$f(\vec{x}) = 0.1146x_1 + 0.1294x_2 + 0.2502x_3 + 0.0608x_4 + 0.1962x_5 + 0.1245x_6 + 0.1241x_7$
3	$f(\vec{x}) = 0.1159x_1 + 0.1303x_2 + 0.2529x_3 + 0.0831x_4 + 0.1675x_5 + 0.1253x_6 + 0.1249x_7$
4	$f(\vec{x}) = 0.1143x_1 + 0.1315x_2 + 0.2533x_3 + 0.1382x_4 + 0.1148x_5 + 0.1240x_6 + 0.1238x_7$
5	$f(\vec{x}) = 0.1144x_1 + 0.1298x_2 + 0.2488x_3 + 0.1256x_4 + 0.1293x_5 + 0.1260x_6 + 0.1261x_7$

Tabla 4.48. Pesos obtenidos en cada iteración y pesos promedio

Iteración	Coefficiente de correlación de Spearman	Prueba de significancia
1	-0.49030	$5.610571178134415 \times 10^{-18}$
2	-0.56673	$2.0000206401633223 \times 10^{-24}$
3	-0.57967	$1.0009009010700406 \times 10^{-25}$
4	-0.55794	$1.419841718513112 \times 10^{-23}$
5	-0.62509	$8.796997117315125 \times 10^{-31}$

Paso 5. Obtención del modelo

De esta manera, la aplicación de la validación cruzada asegura que cada instancia del conjunto de datos participe en los procesos de entrenamiento y prueba, permitiendo así obtener una estimación más robusta, objetiva y representativa del rendimiento real del modelo. Este procedimiento no solo contribuye a reducir el sesgo derivado de una única partición de los datos; sino que también proporciona una evaluación más confiable respecto a su capacidad de generalización en distintos escenarios.

Adicionalmente, al considerar todas las particiones posibles durante el proceso de validación, se garantiza que los coeficientes finales del modelo reflejen un comportamiento promedio, reduciendo la influencia de valores atípicos o estructuras particulares presentes en un subconjunto específico. Como resultado, el modelo multicriterio obtenido se fortalece metodológicamente, configurándose

como una representación más sólida, estable y representativa del fenómeno de estudio.

Finalmente, la ecuación 4.25 presenta el modelo matemático multicriterio obtenido; el cual se define como **“Indicador de Accesibilidad a Servicios (IAS)”**. Este modelo integra de manera ponderada las variables relevantes, proporcionando una herramienta cuantitativa para evaluar la accesibilidad a servicios, específicamente las escuelas de nivel primaria.

$$IAS(\bar{x}) = 0.1089x_1 + 0.1214x_2 + 0.2143x_3 + 0.1335x_4 + 0.1726x_5 + 0.1248x_6 + 0.1245x_7 \quad (4.25)$$

Donde:

- IAS** Modelo de accesibilidad a servicios
- x_1** SBASC, Población de 15 años o más sin educación básica
- x_2** OVHAC, Viviendas con hacinamiento
- x_3** PGRAPROES, Grado promedio de escolaridad
- x_4** PVPH_SNBIEN, Viviendas particulares habitadas sin ningún bien
- x_5** PVPH_SINTIC, Viviendas particulares habitadas sin tecnologías de la información y de la comunicación (TIC)
- x_6** NUM_ESC, Número de escuelas ubicadas en la localidad
- x_7** DIST_PR_PIE, Distancia promedio del origen al destino, en la modalidad de transporte a pie

2.3) Dimensionalidad del modelo multicriterio

La dimensionalidad hace referencia a la variedad de dimensiones y unidades de medida presentes en un modelo. Es decir, describe la estructura del modelo en función de los distintos variables y sus respectivas unidades. Por ejemplo, el criterio de distancia puede medirse en metros, el de tiempo en segundos, y el de escolaridad en porcentaje.

Por otro lado, la no dimensionalidad es el proceso de eliminar las unidades de medida para facilitar un análisis más directo y global del modelo. Este enfoque es particularmente útil en modelos multicriterio, donde se combinan diversas variables con diferentes unidades en una unidad común. Al transformar las variables en valores sin unidades, se simplifica la comparación, evaluación y selección de alternativas; ya que las unidades no interfieren en el análisis, la interpretación ni en la toma de decisiones.

Para convertir un modelo matemático multicriterio en un modelo adimensional (sin unidades de medida), se normalizan los valores de las variables. Cada debe ser dividido entre una variable de referencia correspondiente; la cual está expresada en las mismas unidades de medida. De esta manera, se eliminan las unidades y el resultado es un valor sin dimensiones; es decir, un número puro. La variable de referencia puede ser definida por diversos métodos, tales como la media, la mediana, la moda, el valor máximo, el mínimo, la desviación estándar o la varianza de los valores de las variables.

El primer paso consistió en la identificación de las unidades de medida de cada uno de las variables que conforman el modelo matemático multicriterio obtenido (definido como indicador de accesibilidad a servicios). La Tabla 4.49 se presentan las unidades de medida de las variables que conforman el modelo de la ecuación 4.26.

$$IAS(\bar{x}) = 0.1089x_1 + 0.1214x_2 + 0.2143x_3 + 0.1335x_4 + 0.1726x_5 + 0.1248x_6 + 0.1245x_7 \quad (4.26)$$

Tabla 4.49. Unidades de medida de las variables

Variables	Unidades de medida
SBASC, Población de 15 años o más sin educación básica	Porcentaje
OVHAC, Viviendas con hacinamiento	Porcentaje
PGRAPROES, Grado promedio de escolaridad	Decimales
PVPH_SNBIEN, Viviendas particulares habitadas sin ningún bien	Porcentaje
PVPH_SINTIC, Viviendas particulares habitadas sin tecnologías de la información y de la comunicación (TIC)	Porcentaje
NUM_ESC, Número de escuelas ubicadas en la localidad	Entero
DIST_PR_PIE, Distancia promedio del origen al destino, en la modalidad de transporte a pie	Metros

El siguiente paso consistió en definir las variables de referencia para cada variable. Estas variables están expresadas en las mismas unidades de medida que las variables, lo que asegura su coherencia y facilita la comparación entre ellos. Para el caso de estudio, se consideró el valor medio de las variables, calculado a partir de diversos experimentos previos; lo que garantiza una representación precisa y confiable de los valores en diferentes condiciones. La ecuación 4.27 muestra la definición de las variables de referencia correspondientes a cada variable; mientras que la Tabla 4.50 presenta los valores medios de las variables de referencia obtenidos de dichos experimentos.

$$\begin{aligned}
 x_{1r}: & \text{Valor medio del \% Población de 15 años o más sin educación básica (SBASC)} \\
 x_{2r}: & \text{Valor medio del \% Viviendas con hacinamiento (OVHAC)} \\
 x_{3r}: & \text{Valor medio del Grado promedio de escolaridad (PGRAPROES)} \\
 x_{4r}: & \text{Valor medio del \% Viviendas particulares habitadas sin ningún bien (PVPH_SNBIEN)} \\
 x_{5r}: & \text{Valor medio del \% Viviendas particulares habitadas sin tecnologías de la información y de la comunicación (TIC) (PVPH_SINTIC)} \\
 x_{6r}: & \text{Valor medio del Número de escuelas ubicadas en la localidad (NUM_ESC)} \\
 x_{7r}: & \text{Valor medio de la Distancia promedio del origen al destino, en la modalidad de transporte a pie (DIST_PR_PIE)}
 \end{aligned} \quad (4.27)$$

Tabla 4.50. Unidades y valores de las variables de referencia

Variable de referencia	Valor medio	Unidades de medida
x_{1r}	16.9850	Porcentaje
x_{2r}	19.9648	Porcentaje
x_{3r}	9.0679	Continuo
x_{4r}	0.0352	Porcentaje
x_{5r}	0.0903	Porcentaje
x_{6r}	5	Entero
x_{7r}	5071.3050	Metros

Posteriormente se aplicó el proceso de normalización de los datos; la cual consistió en dividir cada variable entre su variable de referencia correspondiente, para eliminar las unidades asociadas a dichas variables, aplicando las reglas de cancelación de términos comunes. Por ejemplo, la variable x_7 , asociado a DIST_PR_PIE (distancia medida en metros), se dividió entre su variable de referencia x_{7r} . Como resultado, la variable x_{a7} obtenido adquiere un valor puramente numérico, es decir, se convierte en la variable con magnitud adimensional, véase ecuación 4.28.

$$x_{a7} = \frac{x_7 \text{ metros}}{x_{7r} \text{ metros}} \rightarrow x_{a7} = \frac{x_7}{x_{7r}} \quad (4.28)$$

Este procedimiento se aplicó de manera análoga a todas las variables del modelo, transformando cada uno en una versión adimensional mediante la división entre sus respectivas variables de referencia. De esta forma, se eliminaron las unidades asociadas, garantizando que las variables sean comparables entre sí y evitando sesgos derivados de diferencias en las escalas de medida. La ecuación 4.29 muestra el modelo matemático multicriterio adimensional, definido como indicador de accesibilidad a servicios; asegurando la homogeneidad dimensional, la consistencia entre variables y una mejor interpretabilidad de los resultados al reflejar únicamente relaciones relativas sin la influencia de unidades físicas.

Asimismo, se estructuró un árbol de decisión para representar las variables del modelo y sus respectivos pesos. En el modelo multicriterio, el árbol de decisión es una herramienta visual y analítica que permite organizar y evaluar distintas alternativas considerando múltiples variables. En esta estructura, cada nodo representa una decisión o una evaluación basada en una variable específica. En Finalmente se puede observar que la variable de mayor impacto es el grado de escolaridad promedio (PGRAPROES), seguido de las viviendas sin acceso a internet (PVPH_SINTIC) y luego las viviendas que no cuentan con algún bien (PVPH_SNBIEN). Asimismo, se elaboró un árbol de decisión a partir de los valores de los pesos de las variables. En un modelo multicriterio, el árbol de decisión se utiliza para estructurar y analizar las distintas alternativas disponibles considerando varios factores que puedan impactar (Baird, 1978), (Eiselt & Sandblom, 2022). A continuación, se describen los elementos de un árbol de decisión en este contexto del caso de estudio:

- o *Criterios:* cada nodo principal en el árbol representa un criterio fundamental para evaluar la accesibilidad a servicios. Estos criterios agrupan diversos variables que influyen en la toma de decisiones. Criterio de educación está relacionado con el

acceso a servicios educativos. El criterio de vivienda considera características de la infraestructura habitacional. Y el criterio espacial toma en cuenta factores geográficos y de proximidad a los servicios.

- o *Alternativas*: En cada nivel del árbol, se evalúan diferentes opciones o alternativas en función de los criterios establecidos. Las ramas conectan las alternativas, reflejando los posibles escenarios que se derivan de su evaluación.
- o *Evaluación y ponderación*: Cada alternativa se evalúa en función de su de los valores de cada criterio. Las variables y criterios se ponderan según su nivel de importancia, lo que permite determinar su impacto dentro del modelo. La estructura del árbol refleja cómo cada variable aporta a su criterio, estableciendo una jerarquía en la evaluación.

Valor de decisión: El árbol de decisión incorpora una estructura que permite calcular una puntuación global del indicador de accesibilidad a servicios. Cada variable contribuye al desempeño total del modelo, permitiendo una evaluación integral basada en múltiples factores. En este caso, los valores finales de accesibilidad se determinan combinando la ponderación de cada criterio y variable dentro de la estructura del árbol.

Derivado de lo anterior, la Figura 4.42 muestra la estructura del árbol de decisión del modelo multicriterio. Aunado a ello, en la Figura 4.43 se muestran valores de pesos relativos, reflejando su nivel de importancia dentro del modelo de manera ordena (en un ranking).

Indicador de Accesibilidad a Servicios

$$IAS(\bar{x}) = 0.1089x_{a1} + 0.1214x_{a2} + 0.2143x_{a3} + 0.1335x_{a4} + 0.1726x_{a5} + 0.1248x_{a6} + 0.1245x_{a7} \quad (4.29)$$

Donde:

IAS Modelo de accesibilidad a servicios

x_{a1} SBASC, Población de 15 años o más sin educación básica

x_{a2} OVHAC, Viviendas con hacinamiento

x_{a3} PGRAPROES, Grado promedio de escolaridad

x_{a4} PVPH_SNBIEN, Viviendas particulares habitadas sin ningún bien

x_{a5} PVPH_SINTIC, Viviendas particulares habitadas sin tecnologías de la información y de la comunicación (TIC)

x_{a6} NUM_ESC, Número de escuelas ubicadas en la localidad

x_{a7} DIST_PR_PIE, Distancia promedio del origen al destino, en la modalidad de transporte a pie

Donde el Indicador de Accesibilidad a Servicios (IAS) es un modelo matemático que cumple $IAS: \mathbb{R}^7 \rightarrow \mathbb{R}$, con $x_{ai} \in \mathbb{R}$ para $i = \overline{1,7}$, lo que implica que el modelo recibe como entrada un vector de siete variables cuantitativas asociadas con los criterios (dimensiones) de vivienda, educación y espacial; y devuelve un valor escalar que representa el nivel compuesto de accesibilidad. Este valor sintetiza la información multidimensional en una única medida interpretable, facilitando la comparación entre localidades.

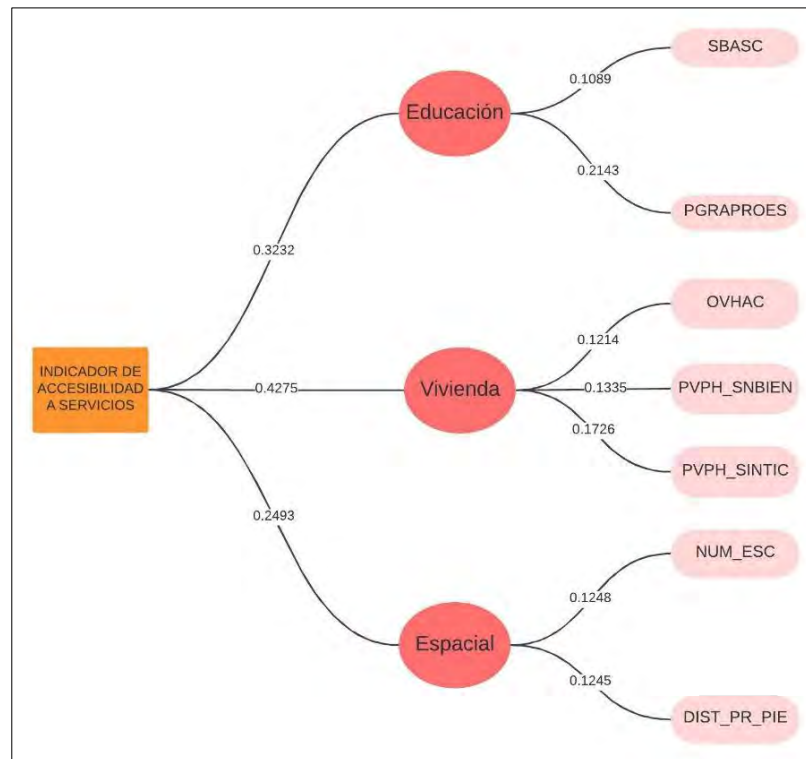


Figura 4.42. Árbol de decisión del modelo multicriterio

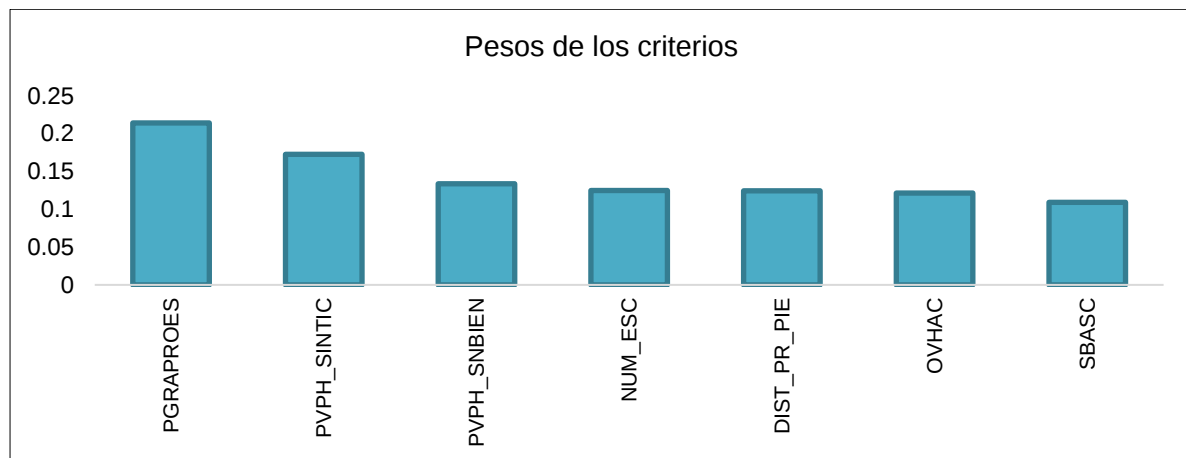


Figura 4.43. Pesos relativos de los criterios en un ranking

4.5. Conclusión

En esta investigación se desarrolló un modelo matemático basado en la Toma de Decisiones Multicriterio para evaluar la accesibilidad a las escuelas de nivel primaria. Este modelo integra factores físicos y sociales, permitiendo cuantificar el nivel de acceso a los servicios educativos y facilitando la toma de decisiones fundamentada.

Se construyó un método de ponderación de variables para asignar los pesos a los coeficientes del modelo de agregación utilizado con un enfoque objetivo que toma la información a partir de los datos; reduciendo la dependencia de la consulta a expertos

y garantizando una evaluación más reproducible y menos subjetiva. El método integró el Análisis de Componentes Principales (PCA) y el Método de Entropía, considerando tanto la variabilidad de los datos como la correlación entre las variables. Una de las principales ventajas de esta combinación de técnicas es que no es menos susceptible a los métodos de transformación de los datos; a diferencia de otros enfoques que llegar a perder la proporcionalidad en la distribución de la información. Esto permite una asignación de pesos más robusta y adaptable.

El enfoque integral implementado para identificar las variables más relevantes permitió una selección objetiva de las variables más significativas, asegurando que el modelo capturara con precisión los factores de mayor impacto en la accesibilidad. Al considerar múltiples perspectivas en la selección de variables, se logró una evaluación más robusta y representativa de la realidad; eliminando redundancias y preservando información clave para el análisis.

El modelo multicriterio obtenido se distingue por su flexibilidad y escalabilidad, ya que puede incorporar nuevas variables sin comprometer su estructura, permitiendo su adaptación a diferentes contextos y niveles de agregación territorial. Su capacidad de ajuste y su enfoque multicriterio lo hacen aplicable a distintas escalas de análisis, desde áreas específicas hasta regiones más amplias; facilitando su integración en procesos de planificación y toma de decisiones en diversos entornos.

Capítulo 5. Validación e interpretación del modelo matemático multicriterio

5.1. Introducción

En este capítulo se presenta la validación del modelo matemático multicriterio desarrollado aplicado al caso de estudio. El proceso de validación considera dos aspectos clave.

El primer aspecto es un análisis comparativo, que consiste en contrastar la coherencia, y concordancia de los resultados obtenidos por el modelo propuesto con los modelos existentes de referencia base. El segundo aspecto corresponde a un análisis de sensibilidad, cuyo objetivo es evaluar la robustez y estabilidad del modelo frente a variaciones en los datos y parámetros. Ambos aspectos proporcionan una validación integral y sólida del modelo desarrollado, garantizando que sea una herramienta confiable y flexible para apoyar la toma de decisiones en la selección de alternativas (Qureshi et al., 1999).

Por otro lado, se realiza la interpretación del modelo multicriterio, denominado Indicador de Accesibilidad a Servicios, con el fin de cuantificar y estructurar el nivel de accesibilidad a los servicios. Esta interpretación es fundamental, ya que facilita la evaluación y comparación entre distintas alternativas; apoyando la toma de decisiones enfocada en la selección de dichas opciones, además ayuda a una mejor comprensión del problema de estudio.

5.1.1. Análisis comparativo

El análisis comparativo es un enfoque utilizado para evaluar la coherencia, y concordancia de los resultados generados por el modelo propuesto, en relación con modelos de referencia o modelos base. La coherencia se refiere a que los resultados del modelo mantengan una lógica interna, sin contradicciones, y reflejen una estructura racional en sus conclusiones. La concordancia implica que dichos resultados estén alineados con la información empírica disponible, es decir, que reproduzcan adecuadamente los patrones y tendencias observadas en los datos reales. Este tipo de evaluación busca identificar patrones comunes; lo que contribuye a fortalecer la aplicabilidad del modelo en situaciones reales, aumentar la confianza en sus resultados y facilitar la realización de ajustes cuando se identifican desviaciones relevantes respecto a las tendencias esperadas (CONAPO CentroGeo, 2022), (Roy & Kar, 2022).

Una de las herramientas empleadas para evaluar la concordancia es el coeficiente de correlación, el cual permite medir el grado de asociación entre los resultados obtenidos por dos modelos (Zafri et al., 2021), (Bucherie et al., 2022). El coeficiente de correlación resulta útil para identificar tendencias comunes, validar la estructura del modelo y verificar su alineación con enfoques consolidados en la literatura.

Asimismo, para evaluar la coherencia del modelo, se utiliza también el coeficiente de correlación de Spearman como herramienta principal, ya que permite examinar la existencia de relaciones lógicas esperadas entre variables fundamentales del sistema analizado. En particular, se analiza si dichas relaciones son directas o inversas. Esto

contribuye a confirmar que el comportamiento del modelo es consistente con los principios teóricos y empíricos que sustentan su construcción.

Los modelos base de referencia son: el Índice de Marginación por localidad, el Indicador de Oportunidades Acumuladas y el Indicador de Accesibilidad Basado en la Distancia. La elección de estos modelos se fundamenta en que todos permiten realizar el análisis a un nivel de agregación por localidad, lo cual es coherente con la escala de estudio. El Indicador de Oportunidades Acumuladas y el Indicador de Accesibilidad Basado en la Distancia, abordan la accesibilidad desde la dimensión espacial; mientras que el Índice de Marginación permite conocer características de las localidades desde una perspectiva social, ofreciendo así una visión complementaria y multidimensional del fenómeno analizado.

Los modelos base de referencia son: el Índice de Marginación por localidad, el Indicador de Oportunidades Acumuladas y el Indicador de Accesibilidad Basado en la Distancia. La elección de estos modelos se fundamenta en que todos permiten realizar el análisis a un nivel de agregación por localidad, lo cual es coherente con la escala de estudio. Por su parte, el Índice de Marginación integra múltiples dimensiones enfocadas en las carencias sociales; lo que permite caracterizar las condiciones estructurales de cada localidad. De esta forma, la combinación de estos modelos ofrece una visión complementaria y multidimensional del fenómeno analizado, al integrar tanto la dimensión espacial de la accesibilidad como las condiciones sociales subyacentes de las localidades.

A continuación, se presentan las expresiones matemáticas correspondientes a cada uno de estos modelos, junto con la perspectiva de análisis que adoptan y las variables consideradas. Todos los modelos están estructurados en cinco niveles, numerados del 1 al 5. En el caso del Índice de Marginación, los valores bajos se asignan al nivel 1 (muy bajo), lo que representa condiciones sociales más favorables y menor marginación. En cambio, los valores altos corresponden al nivel 5 (muy alto), indicando mayores carencias y un mayor rezago social. Por otro lado, en los indicadores de accesibilidad, los valores altos se asignan al nivel 1 (muy bajo), lo que indica una accesibilidad limitada a los servicios; mientras que los valores bajos se clasifican en el nivel 5 (muy alto), representando una mayor accesibilidad, véase Tabla 5.1.

De esta forma, la categorización de los niveles en los modelos de accesibilidad y en el Índice de Marginación tiene un significado opuesto. Mientras que un nivel alto en marginación refleja desventaja social, un nivel alto en accesibilidad representa una condición favorable. Esta relación entre modelos es útil para identificar patrones complementarios o contrastantes entre marginación y accesibilidad, permitiendo un análisis integral de las condiciones territoriales en alguna zona geográfica.

Tabla 5.1. Descripción de los modelos de referencia

Modelos de referencia
Índice de marginación por localidad - Descripción: Mide las carencias de una localidad en términos de variables de población, educación, seguridad social, vivienda y servicios. (CONAPO, 2022) - Variables: Población analfabeta de 15 años, Población de 15 años o más sin educación básica Viviendas sin drenaje ni excusado Viviendas sin energía eléctrica, Viviendas sin agua entubada, Viviendas con piso de tierra, Viviendas con hacinamiento Viviendas sin refrigerador $DP_2 = \sum_{i=1}^n \frac{d_{ij}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2)$ $d_{ij} = x_{rj} - x_{*j}$ es la distancia de la j-ésima variable de la localidad r. σ_j: es la desviación estándar de la variable j, $R_{j,j-1,\dots,1}^2$: es el coeficiente de determinación de la regresión del indicador parcial j con respecto a los otros indicadores ($j-1, j-2, \dots, 1$). $(1 - R_{j,j-1,\dots,1}^2)$: es el factor corrector que evita la duplicidad, al eliminar la información parcial de los indicadores.
Indicador de accesibilidad basado en la distancia - Descripción: Mide el nivel de accesibilidad geográfica de la localidad, basado en la proximidad (Shahid et al., 2009). - Variables: Distancia (Dimensión espacial) $IA = \frac{1}{n} \sum_{j=1}^n D_{ij}$ Donde D_{ij} es la entre la localidad i y la escuela j.
Indicador de accesibilidad de oportunidades acumuladas - Descripción: Mide el nivel de accesibilidad geográfica de la localidad, basado en la disponibilidad de los servicios (Kelobonye et al., 2020). - Variables: Número de servicios (Dimensión espacial) $IA = \sum_{j=1}^n o_j$ Donde o_j es número de escuelas en cada localidad.

El análisis comparativo, se empleó al conjunto de datos de prueba previamente definido en la sección 4.4.2; el cual incluyó 454 instancias, correspondientes al 25 % del total de datos. El procedimiento consistió en evaluar este conjunto el modelo matemático multicriterio propuesto (Indicador de Accesibilidad a Servicios obtenido en la ecuación 4.29), para generar un puntaje global de accesibilidad. Posteriormente, dichos valores fueron clasificados mediante el método de optimización de Jenks.

Los resultados se agruparon en cinco categorías, siguiendo el esquema utilizado por fuentes oficiales como el Índice de Desarrollo Social de la Ciudad de México, la Medición de la Pobreza y el Índice de Rezago Social, entre otros (EVALÚA, 2024), (CONEVAL, 2024), (CONEVAL, 2023).

Con el objetivo de evaluar la concordancia y consistencia del Indicador de Accesibilidad a Servicios propuesto (modelo matemático de accesibilidad multicriterio), se calculó el coeficiente de correlación de Spearman respecto a tres

modelos de referencia. Este análisis permitió identificar el grado de asociación entre el modelo propuesto y cada uno de los modelos de referencia, tanto desde la dimensión social como espacial.

Los resultados del análisis de concordancia y consistencia, obtenidos mediante el coeficiente de correlación de Spearman, evidencian una relación significativa entre el modelo de accesibilidad multicriterio propuesto y los tres modelos de referencia. En particular, se observó una correlación positiva moderada con el Indicador de Oportunidades Acumuladas ($\rho = 0.4433$) y con el Indicador de Accesibilidad basado en la Distancia ($\rho = 0.4716$). Estos resultados reflejan un comportamiento coherente entre los modelos; ya que, a mayor nivel de accesibilidad estimado por el modelo multicriterio, también se incrementan los valores de accesibilidad reportados por los modelos de referencia.

Por otro lado, se identificó una correlación negativa moderada entre el modelo propuesto y el Índice de Marginación por localidad ($\rho = -0.5843$). Este resultado indica que, a mayor marginación, la accesibilidad a servicios tiende a disminuir; lo cual es consistente con lo esperado, dado que mayores niveles de marginación implican carencias sociales y económicas asociadas comúnmente con un acceso más limitado a servicios básicos. La Tabla 5.2 muestra los resultados del coeficiente de correlación

Tabla 5.2. Resultados del análisis de concordancia y consistencia del modelo propuesto

Índice de marginación por localidad $\rho = -0.5843$ p – valor = 7.9890e – 43
Indicador de accesibilidad de Oportunidades acumuladas $\rho = 0.4433$ p – valor = 3.1313e – 23
Indicador de accesibilidad basado en distancia $\rho = 0.4716$ p – valor = 1.8537e – 26

Para reforzar el análisis de concordancia y coherencia, se realizó una representación gráfica que ilustran el comportamiento del modelo matemático multicriterio propuesto en comparación con los modelos de referencia. Esta representación gráfica se realizó a nivel de agregación de localidad y municipal. Estas visualizaciones permiten identificar patrones generales y similitudes espaciales en la distribución de los indicadores. En particular, las gráficas a escala municipal facilitaron una mejor comprensión de las tendencias agregadas, destacando comportamientos consistentes entre los distintos enfoques.

Para complementar el análisis de concordancia y coherencia, se elaboró una representación gráfica que ilustra el comportamiento del modelo matemático multicriterio propuesto en comparación con los modelos de referencia. Esta visualización se desarrolló tanto a nivel de localidad como a nivel municipal, lo que permitió identificar con mayor claridad los patrones de distribución y las similitudes

espaciales entre los indicadores analizados. Las Figura 5.1, Figura 5.1, Figura 5.2, Figura 5.3 y Figura 5.4 muestra el comportamiento a nivel localidad. Por otra parte, las Figura 5.5, Figura 5.6, Figura 5.7 y Figura 5.8 a nivel municipal. En particular, las gráficas a escala municipal muestran que este nivel de agregación facilita una interpretación más clara y precisa del comportamiento entre los modelos; al resaltar de manera más evidente la correspondencia y coherencia entre los distintos enfoques analizados.

Adicionalmente, se representó gráficamente, mediante un mapa de coropletas, el comportamiento del modelo matemático multicriterio de accesibilidad propuesto y el del índice de marginación. Esta visualización cartográfica evidencia una tendencia general: las localidades con mayores niveles de accesibilidad tienden a registrar niveles más bajos de marginación, mientras que aquellas con accesibilidad limitada suelen presentar altos niveles de marginación. Esta relación inversa es coherente con los resultados obtenidos en el análisis de correlación, lo que refuerza la validez del modelo propuesto. La Figura 5.9 muestra esta correspondencia inversa a través de un caso específico en el municipio de Tlalpan y sus localidades, donde el patrón descrito se manifiesta de forma clara y consistente.

Como conclusión del análisis comparativo, a través del análisis de concordancia y coherencia, se puede afirmar que los resultados estadísticos, reflejados en los coeficientes de correlación, junto con el comportamiento observado de manera gráfica, respaldan la validez y consistencia del modelo matemático multicriterio propuesto con metodologías consolidadas en la literatura. La tendencia observada entre del modelo desarrollado refuerza su coherencia frente a otros enfoques reconocidos, lo que consolida su pertinencia como herramienta analítica robusta para el estudio de la accesibilidad.

Como conclusión del análisis comparativo se puede afirmar que los resultados estadísticos, reflejados en los coeficientes de correlación, y en las gráficas de comportamiento, respaldan la validez y consistencia del modelo matemático multicriterio propuesto, con respecto de metodologías consolidadas en la literatura. La tendencia observada en el modelo desarrollado refuerza su coherencia y concordancia frente a otros enfoques reconocidos, consolidando su pertinencia como herramienta analítica robusta para el estudio de la accesibilidad.

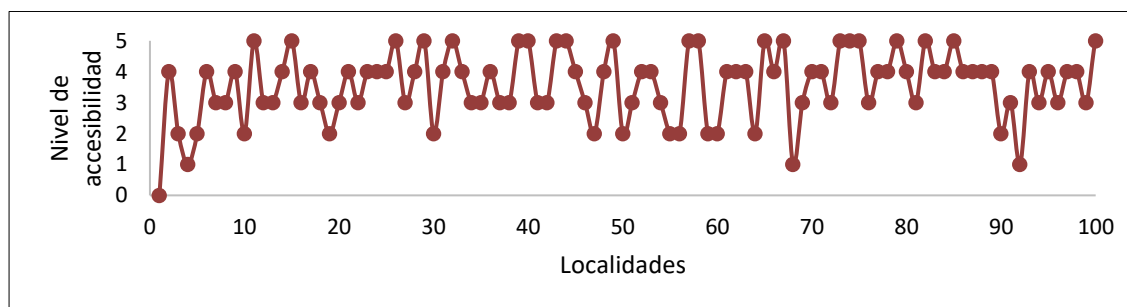


Figura 5.1. Modelo multicriterio de accesibilidad a servicios por localidad propuesto

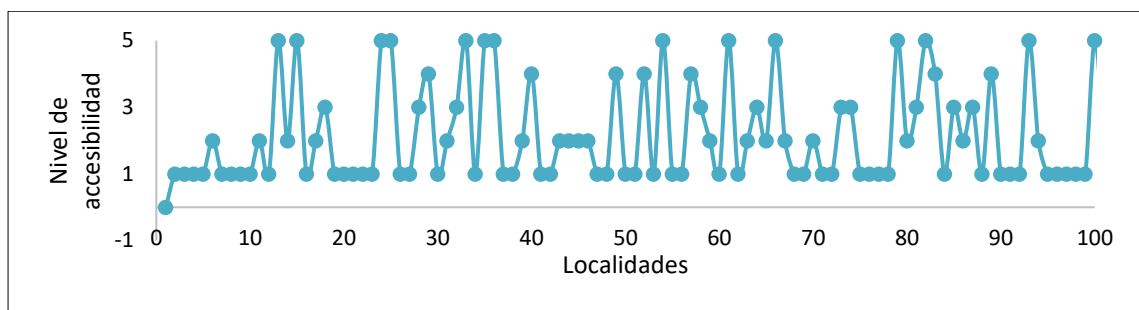


Figura 5.2. Indicador de oportunidades acumuladas por localidad

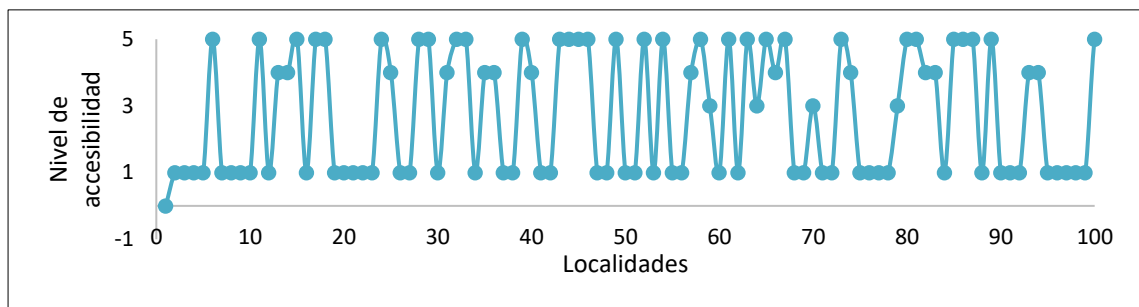


Figura 5.3. Indicador basado en la distancia por localidad

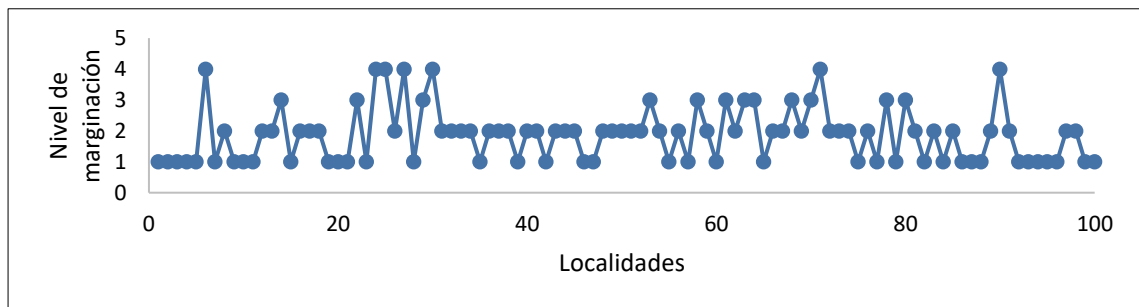


Figura 5.4. Índice de marginación por localidad

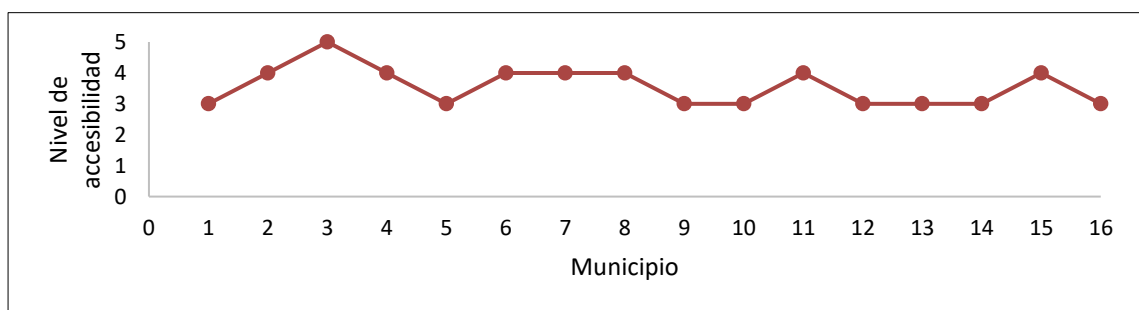


Figura 5.5. Modelo multicriterio de accesibilidad a servicios por municipio propuesto

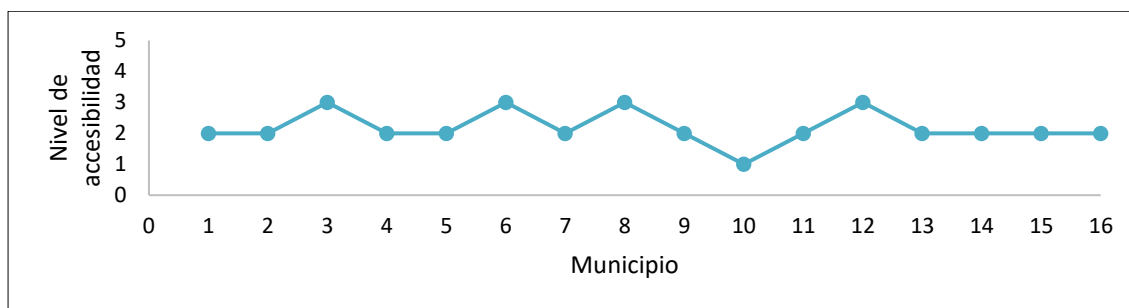


Figura 5.6. Indicador de acumuladas por municipio

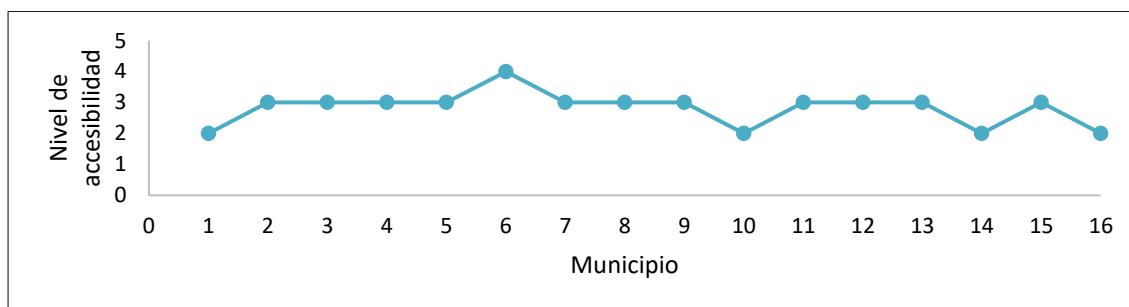


Figura 5.7. Indicador basado en la distancia por municipio

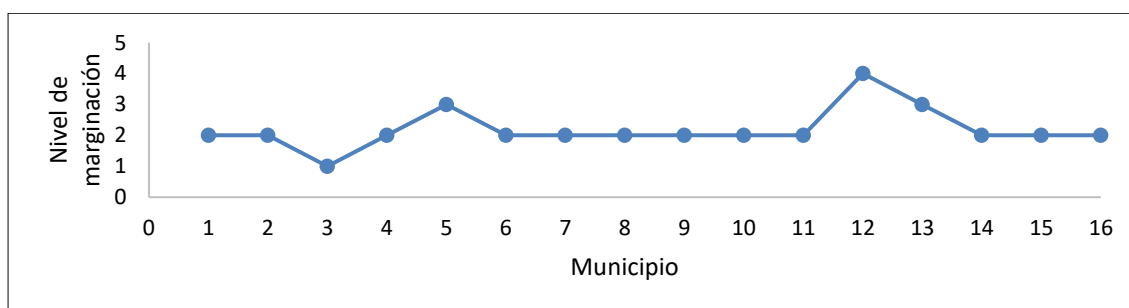


Figura 5.8. Índice de marginación por localidad

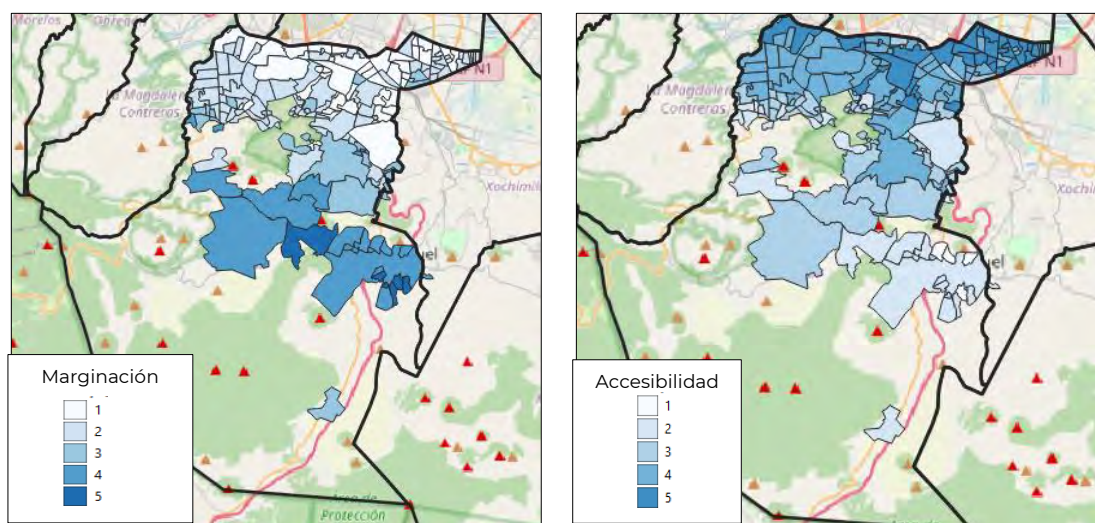


Figura 5.9. Comparación del modelo de accesibilidad propuesto y el índice de marginación

5.1.2. Análisis de sensibilidad

El análisis de sensibilidad es una etapa dentro de la metodología de la Toma de Decisiones Multicriterio y ampliamente utilizada en la literatura; ya que constituye la fase final del proceso. Este análisis consiste en introducir pequeñas variaciones en los datos del modelo para evaluar su impacto en la solución. Dado que los modelos de decisión suelen ser complejos y presentan cierto grado de incertidumbre; este análisis permite explorar distintos escenarios, proporcionando una mejor comprensión de la estabilidad, confiabilidad y robustez del modelo. (Ishizaka & Nemery, 2013), (Muñoz, et al, 2016), (Indriyantho et al., 2024).

Este enfoque permite explorar distintos escenarios hipotéticos mediante modificaciones en la información disponible, evaluando su impacto en la solución final. Esto facilita la identificación de los factores más influyentes en el modelo y la validación de su solidez ante diversas condiciones. El análisis de sensibilidad permite realizar evaluaciones del tipo "qué pasaría si", proporcionando una visión más completa sobre la estabilidad y robustez del modelo. En términos metodológicos, este análisis resulta fundamental para examinar la consistencia del modelo, ya que permite observar cómo responde ante posibles variaciones, incertidumbres o perturbaciones en los datos de entrada y/o en los parámetros considerados. Esta capacidad de adaptación frente a escenarios alternativos fortalece la confiabilidad del modelo propuesto. A través de este proceso, se puede examinar el efecto de cambios en la ponderación de los criterios, ajustes en los datos de entrada o la exclusión de ciertos factores dentro del proceso de decisión, entre otros aspectos (Narváez, 2021), (Gigović et al., 2017), (Erbaş et al., 2018), (Fuentes et al., 2020), (Bierman et al., 1996).

El análisis de sensibilidad se clasifica en métodos locales y globales. Los métodos locales evalúan el impacto de variaciones en parámetros específicos, manteniendo constantes los demás factores. En contraste, los métodos globales permiten que todos los factores de entrada varíen dentro de un rango de incertidumbre, proporcionando una visión más amplia del comportamiento del modelo (Malczewski, 2017).

Dado que el análisis de sensibilidad tiene como objetivo evaluar la estabilidad y robustez del modelo frente a variaciones en sus componentes, en este estudio se implementó un enfoque integral que contempla diferentes perspectivas metodológicas. En particular, se consideraron diversos enfoques de análisis de sensibilidad con el propósito de explorar de manera completa el comportamiento del modelo ante posibles cambios en los datos o parámetros. A continuación, se describen los enfoques utilizados.

- o *Análisis de sensibilidad basado en técnicas no paramétricas:* Se emplearon métodos no paramétricos para generar valores de las variables que replicaran el comportamiento observado en los datos reales. Estos valores simulados fueron evaluados en el modelo con el propósito de examinar su estabilidad frente a escenarios que reflejan dinámicas similares a las del fenómeno observado.
- o *Análisis de sensibilidad basado en simulación estadística:* Se generaron valores aleatorios para las variables utilizando una distribución uniforme, con el objetivo de simular hipotéticos cambios en su comportamiento. Estos valores fueron

posteriormente evaluados en el modelo para analizar su estabilidad ante variaciones en los datos de entrada.

- o *Análisis de sensibilidad basado en cambios en los pesos de las variables:* Se modificaron de manera aleatoria los pesos asignados a las variables dentro del modelo, utilizando un enfoque de tipo cognitivo; con el fin de evaluar el impacto de dichas variaciones sobre los resultados finales y determinar la consistencia del modelo frente a diferentes esquemas de ponderación.

En conjunto, estos tres enfoques permitieron realizar un análisis de sensibilidad integral y representativo, considerando distintos orígenes de incertidumbre: aleatoria (simulación), estructural (no paramétrico) y cognitiva (preferencias). De esta manera, se buscó garantizar que los resultados del modelo multicriterio fueran consistentes, estables y aplicables a una diversidad de escenarios posibles dentro del contexto de análisis.

Análisis de sensibilidad basado en técnicas no paramétricas

En este caso, se generaron datos sintéticos que reproducen el comportamiento estadístico de los datos originales con el propósito de analizar la estabilidad de los resultados de forma más flexible y robusta ante diferentes escenarios. La generación de estos datos permite evaluar cómo se comportan el modelo cuando se introducen pequeñas variaciones en los datos de entrada, sin depender exclusivamente de los datos reales. Esta estrategia resulta especialmente útil en contextos donde se requiere validar la consistencia de los resultados, explorar posibles escenarios futuros o simular condiciones que aún no se han observado en los datos originales. Así, se amplían las posibilidades de análisis y se mejora la capacidad de adaptación del estudio a distintos contextos o incertidumbres en la información disponible.

Para este análisis, se generaron datos sintéticos que conservan el comportamiento estadístico de los datos originales, dado que las variables no siguen una distribución normal. Para ello, se empleó la técnica no paramétrica conocida como Estimación de Densidad por Kernel (KDE, por sus siglas en inglés); la cual permite simular datos manteniendo el patrón estadístico observado sin necesidad de asumir una forma específica de distribución.

El proceso consistió en aplicar KDE para estimar la distribución de las variables originales y generar datos sintéticos coherentes con dicha distribución. Las Figuras 5.10 a 5.12 presentan histogramas comparativos de los datos reales y sintéticos para cada variable; ofreciendo una representación gráfica clara de sus distribuciones y comportamiento de los datos. Asimismo, las Tablas 5.4 y 5.5 resumen estadísticamente los datos originales y los generados, respectivamente. Los cálculos fueron realizados en Python utilizando la librería NumPy para el manejo de los datos; `scipy.stats.gaussian_kde` para generar los datos sintéticos mediante KDE. La visualización de los resultados se llevó a cabo con la librería matplotlib.

La Estimación de Densidad por Kernel es un método utilizado para calcular la densidad de probabilidad de una variable aleatoria de manera no paramétrica. Este procedimiento emplea una función base, denominada kernel, que suaviza los puntos de datos observados y genera una estimación continua de la distribución. Esto

permite una visualización flexible y detallada de la forma de la distribución, siendo especialmente útil cuando no se dispone de conocimiento previo sobre su naturaleza (Jurado, e al, 2023).

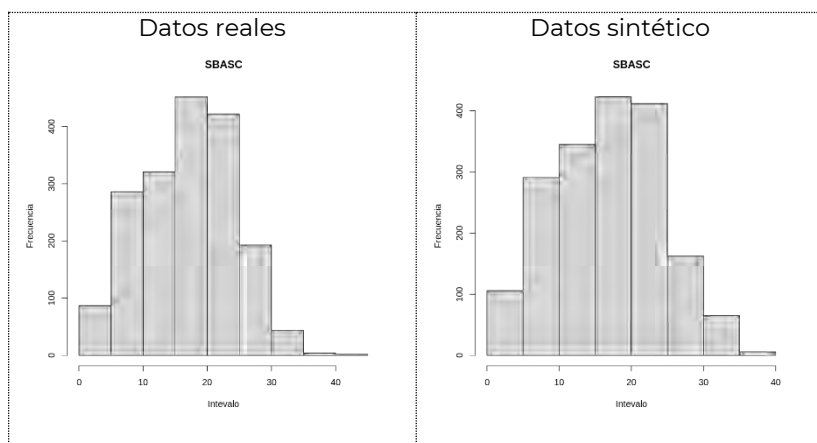


Figura 5.10. Histogramas SBASC: datos reales y simulados

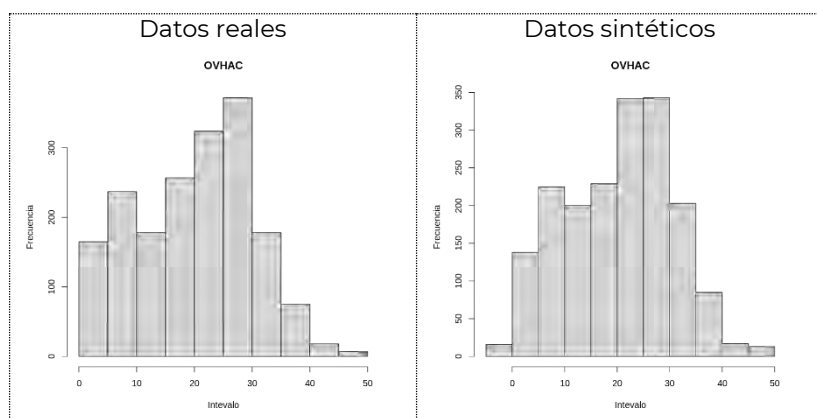


Figura 5.11. Histogramas de OVHAC: datos reales y simulados

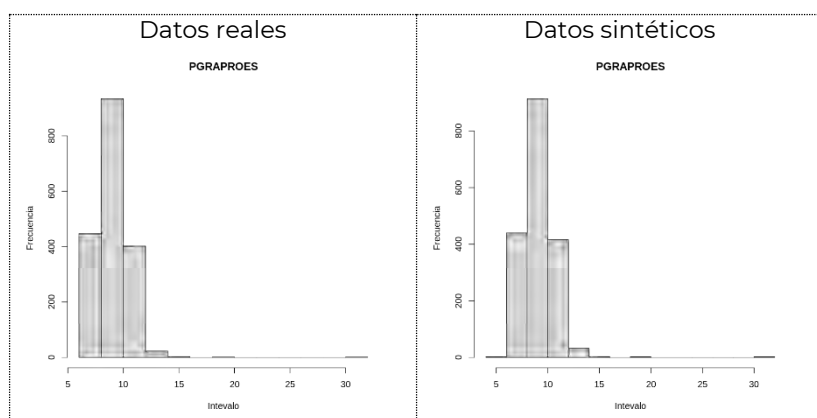


Figura 5.12. Histogramas de PGRAPROES: datos reales y simulados

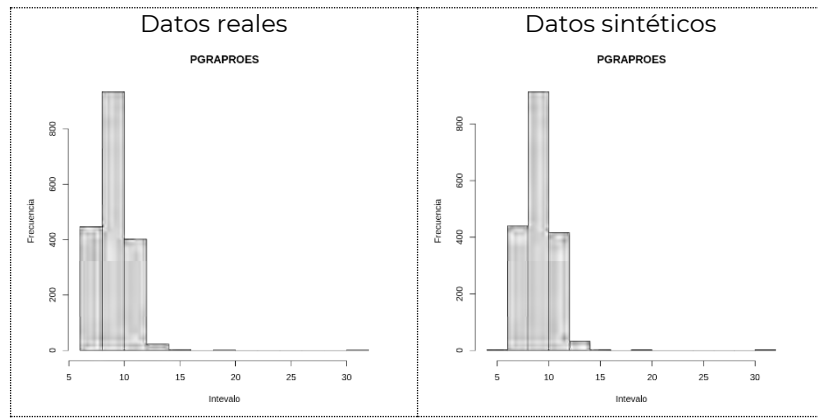


Figura 5.13. Histogramas de PVPH_SNBIEN: datos reales y simulados

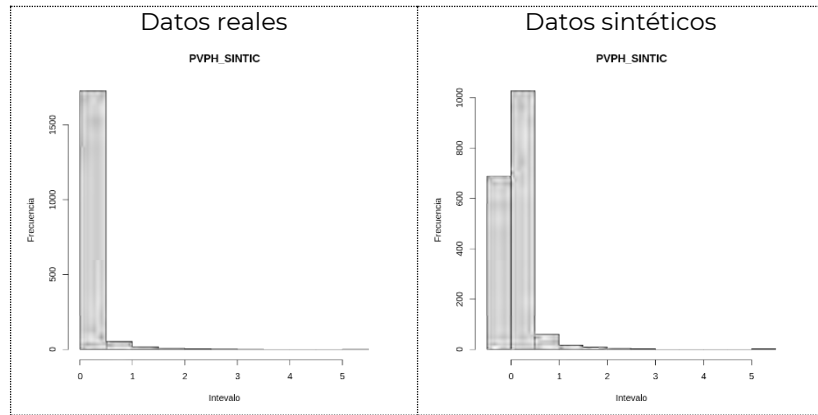


Figura 5.14. Histogramas de PVPH_SINTIC: datos reales y simulados

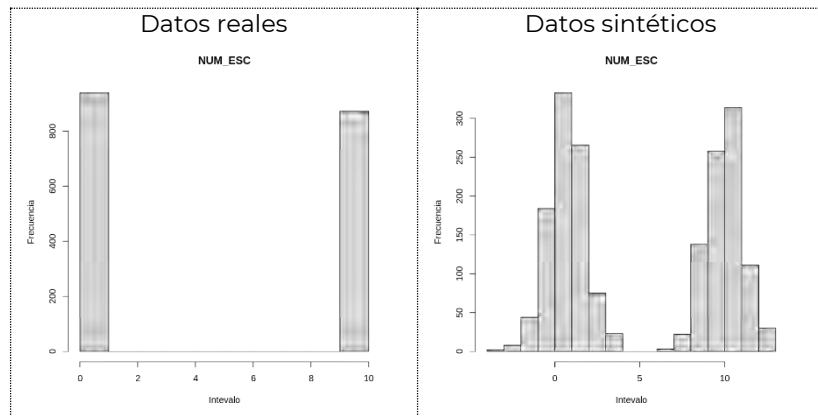


Figura 5.15. Histogramas de NUM_ESC: datos reales y simulados

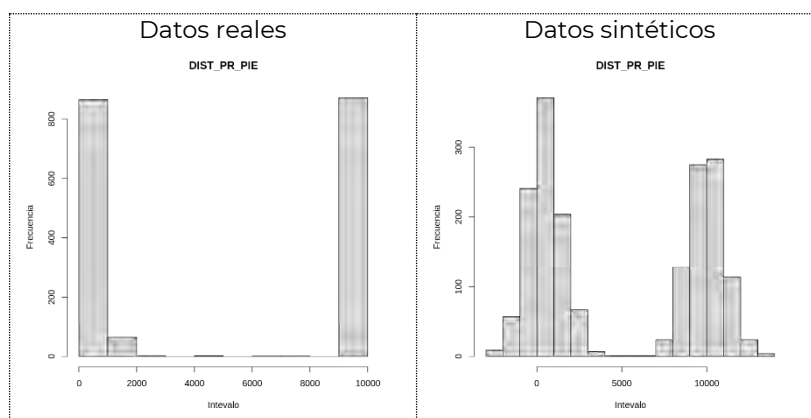


Figura 5.16. Histogramas de SBASC: datos reales y simulados

Tabla 5.3. Descripción estadística de las variables originales

	SBASC	OVHAC	PGRAPROES	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
Total de registros	1811	1811	1811	1811	1811	1811	1811
Valores nulos	0	0	0	1583	1378	0	1
Datos inconsistentes	0	0	0	0	0	0	0
Mínimo	2.163	0.411	6.221	0.000	0.000	0.067	0
Máximo	44.921	47.295	31.250	2.421	5.306	10.000	10000.000
Rango	42.758	46.884	25.029	2.421	5.306	9.933	10000.000
Suma	30759.869	36156.214	16421.898	63.688	163.516	9315.824	9184133.841
Mediana	17.479	21.378	9.061	0	0	1	1275.000
Media	16.985	19.965	9.068	0.035	0.090	5.144	5071.305
SE.mean	0.170	0.236	0.034	0.003	0.007	0.110	111.952
CI.mean.0.95	0.334	0.463	0.067	0.007	0.013	0.216	219.568
Varianza	52.426	101.080	2.123	0.021	0.080	21.969	22697554.517
Desviación estándar	7.241	10.054	1.457	0.146	0.282	4.687	4764.195
Coefficiente de variación	0.426	0.504	0.161	4.148	3.124	0.911	0.939

Tabla 5.4. Descripción estadística de las variables simuladas

	SBASC	OVHAC	PGRAPROES	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
Total de registros	1811	1811	1811	1811	1811	1811	1811
Valores nulos	0	0	0	0	0	0	0
Datos inconsistentes	0	0	0	0	0	0	0
Mínimo	0.388	-2.895	5.865	-0.103	-0.272	-3.520	-2649.559
Máximo	38.476	48.603	31.967	2.454	5.384	12.899	13272.177
Rango	38.088	51.498	26.103	2.557	5.656	16.419	15921.736
Suma	30347.303	36834.713	16496.448	61.271	171.128	9389.993	8984411.360
Mediana	17.087	21.529	9.099	0.004	0.026	2.847	2177.004
Media	16.757	20.339	9.109	0.034	0.094	5.185	4961.022
SE.mean	0.175	0.242	0.038	0.003	0.007	0.112	114.343
CI.mean.0.95	0.344	0.475	0.074	0.007	0.015	0.220	224.257
Varianza	55.638	106.051	2.583	0.022	0.102	22.772	23677444.457
Desviación estándar	7.459	10.298	1.607	0.148	0.319	4.772	4865.947
Coefficiente de variación	0.445	0.506	0.176	4.376	3.372	0.920	0.981

Una vez generados los datos sintéticos de las variables mediante KDE, los valores de cada variable en el conjunto de datos original fueron reemplazados por los datos sintéticos generados, manteniendo sin cambios el resto de las variables en el conjunto. Esto resultó en la creación de siete conjuntos de datos distintos (siete escenarios diferentes). Cada uno de estos conjuntos fue evaluado utilizando el modelo multicriterio denominado Indicador de Accesibilidad a Servicios. Para clasificar las alternativas, se aplicó el método de optimización de intervalos de Jenks, que dividió los datos en cinco categorías.

Finalmente, los resultados del análisis de sensibilidad, basado en técnicas no paramétricas, fueron evaluados mediante los coeficientes de correlación de Spearman y Kendall; con el fin de medir la consistencia entre los valores originales y los datos generados mediante KDE. Los resultados obtenidos confirmaron que el modelo mantiene su estabilidad frente a variaciones en los datos de entrada, mostrando un comportamiento similar al de los valores originales. Los detalles completos de los resultados se presentan en la Tabla 5.5. Este análisis permitió realizar una experimentación sistemática, garantizando que los valores generados representaran un rango realista de posibles escenarios, véase Figura 5.17.

Tabla 5.5 Resultados de las métricas del análisis de sensibilidad de técnicas no paramétricas

Escenarios Cambios en los datos de entrada Variables	Coeficiente de correlación Spearman	Coeficiente de correlación Kendall
SBASC	0.8831	0.8375
OVHAC	0.7853	0.7228
PGRAPROES	0.8640	0.8058
PVPH_SNBIEN	0.8482	0.7970
PVPH_SINTIC	0.8224	0.7637
NUM_ESC	0.8224	0.7652
DIST_PR_PIE	0.8831	0.8375

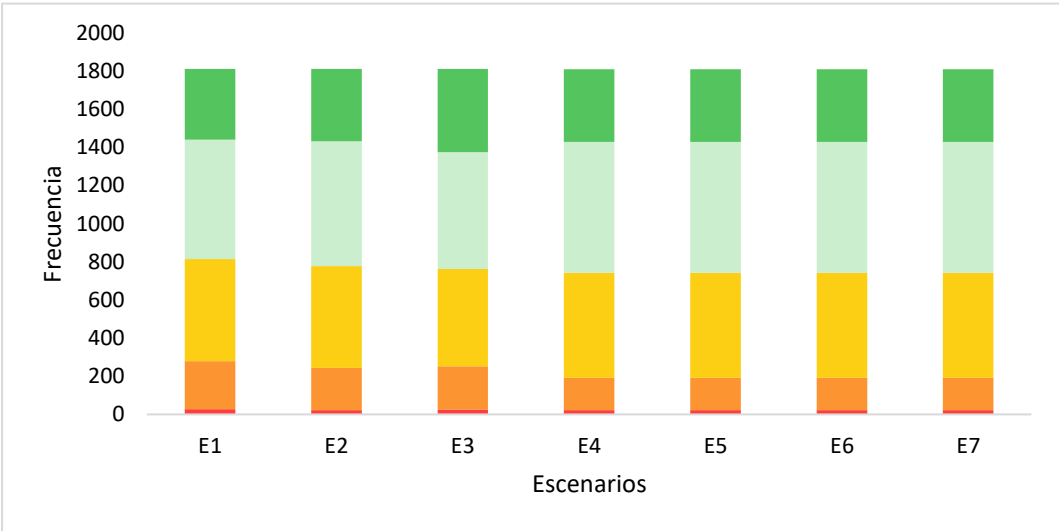


Figura 5.17. Distribución por categorías del análisis con técnicas no paramétricas

Análisis de sensibilidad basado en simulación estadística

En este caso, se generaron datos con un comportamiento correspondiente a una distribución uniforme con el objetivo de evaluar la variabilidad y la robustez del modelo frente a fluctuaciones aleatorias en los datos de entrada. La distribución uniforme permite simular escenarios en los que todos los valores dentro de un rango determinado tienen igual probabilidad de ocurrencia; lo cual resulta útil para probar cómo responde el modelo ante situaciones de incertidumbre sin sesgo hacia valores específicos. Al incorporar esta variabilidad aleatoria controlada, se puede analizar la estabilidad y la sensibilidad del modelo, identificando posibles rangos de comportamiento y garantizando que los resultados sean consistentes incluso bajo condiciones de datos impredecibles o fluctuantes. Esto contribuye a fortalecer la confiabilidad del análisis y la capacidad del modelo para adaptarse a diferentes condiciones y escenarios reales.

El proceso consistió en introducir perturbaciones en los datos originales, enfocándose en las variables del modelo. Para ello, se generaron valores sintéticos utilizando una distribución uniforme continua. Esta distribución permite explorar variaciones amplias en los datos sin introducir sesgos asociados a distribuciones paramétricas específicas (García et al., 2013). El método matemático utilizado para generar estos valores se basa en el muestreo aleatorio uniforme, que se describe con la fórmula 5.1.

$$x = a + (b - a)U \quad (5.1)$$

donde a y b son los límites inferior y superior del intervalo, y U es un número aleatorio generado a partir de una distribución uniforme estándar en el intervalo $[0, 1]$.

Los parámetros de la distribución se establecieron con base en los valores mínimos y máximos observados en las variables de los datos originales, asegurando que los valores generados representaran un rango realista de posibles escenarios. Los cálculos fueron realizados mediante el lenguaje de programación Python con la librería NumPy. La Figura 5.18, Figura 5.19 y Figura 5.20 presentan los histogramas de los datos generados sintéticos para cada variable, proporcionando una visualización de su distribución. Asimismo, la Tabla 5.6 muestra un resumen estadístico de los datos.

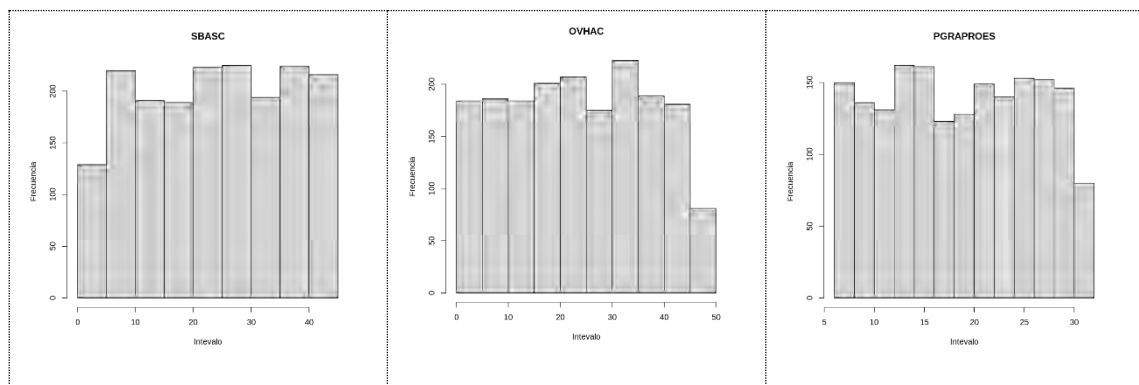


Figura 5.18. Histogramas de las variables con distribución uniforme. Parte 1

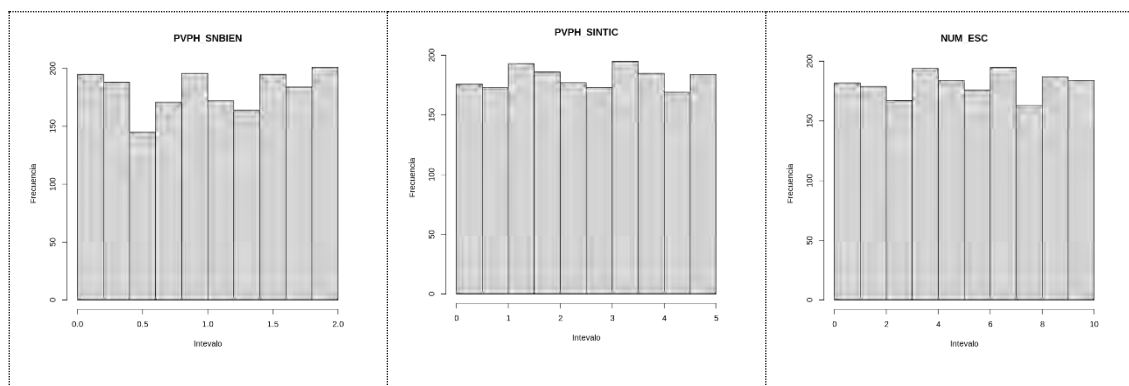


Figura 5.19. Histogramas de las variables con distribución uniforme. Parte 2

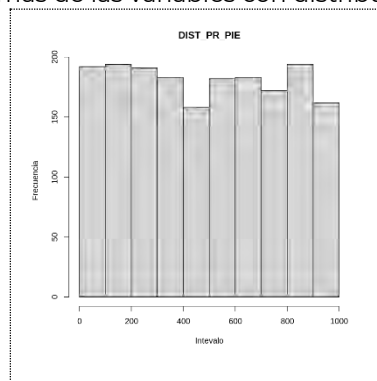


Figura 5.20. Histogramas de las variables con distribución uniforme. Parte 3

Tabla 5.6 Descripción estadística de las variables con distribución uniforme

	SBASC	OVHAC	PGRAPROES	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
nbr.val	1811	1811	1811	1811	1811	1811	1811
nbr.null	0	0	0	0	0	0	0
nbr.na	0	0	0	0	0	0	0
min	2.001	0.048	6.014	0.000	0.009	0.011	1.452
max	44.982	46.984	30.992	1.998	4.999	9.999	999.753
range	42.981	46.936	24.978	1.997	4.991	9.988	998.301
sum	42869.295	42998.451	33732.772	1831.551	4546.295	9071.371	889506.001
median	23.999	23.437	18.743	1.008	2.500	4.999	492.424
mean	23.672	23.743	18.627	1.011	2.510	5.009	491.168
SE.mean	0.293	0.316	0.171	0.014	0.034	0.068	6.844
CI.mean.0.95	0.575	0.619	0.335	0.027	0.066	0.133	13.423
var	155.583	180.681	52.985	0.347	2.074	8.313	84831.787
std.dev	12.473	13.442	7.279	0.589	1.440	2.883	291.259
coef.var	0.527	0.566	0.391	0.583	0.574	0.576	0.593

El siguiente paso consistió en reemplazar cada variable original por los correspondientes datos generados de manera sintética, manteniendo el resto del conjunto de datos sin modificaciones. Este proceso se repitió para cada una de las siete variables, generando así siete conjuntos de datos distintos (7 escenarios diferentes). Cada conjunto de datos fue evaluado en el modelo multicriterio, definido como Indicador de Accesibilidad a Servicios. Para la clasificación de alternativas, se aplicó el método de optimización de intervalos de Jenks, el cual clasificó los datos en cinco categorías.

Finalmente, los resultados del análisis de sensibilidad basada en simulación estadística, fueron analizados mediante los coeficientes de correlación de Spearman y Kendall, lo que permitió evaluar la consistencia entre los valores originales y los datos modificados. Los resultados obtenidos evidencian la estabilidad del modelo frente a modificaciones en las variables, que tienen un comportamiento distinto a los datos reales. Los detalles de los resultados se presentan en la Tabla 5.7 y la Figura 5.21. Este análisis permitió la experimentación sistemática con distintas configuraciones de datos.

Tabla 5.7 Resultados de las métricas análisis de sensibilidad en simulación estadística

Escenarios Cambios en los datos de entrada Variables	Coeficiente de correlación Spearman	Coeficiente de correlación Kendall
SBASC	0.8764	0.8345
OVHAC	0.8683	0.8243
PGRAPROES	0.7512	0.6805
PVPH_SNBIEN	0.8554	0.7962
PVPH_SINTIC	0.8361	0.7717
NUM_ESC	0.8123	0.7577
DIST_PR_PIE	0.8136	0.7578

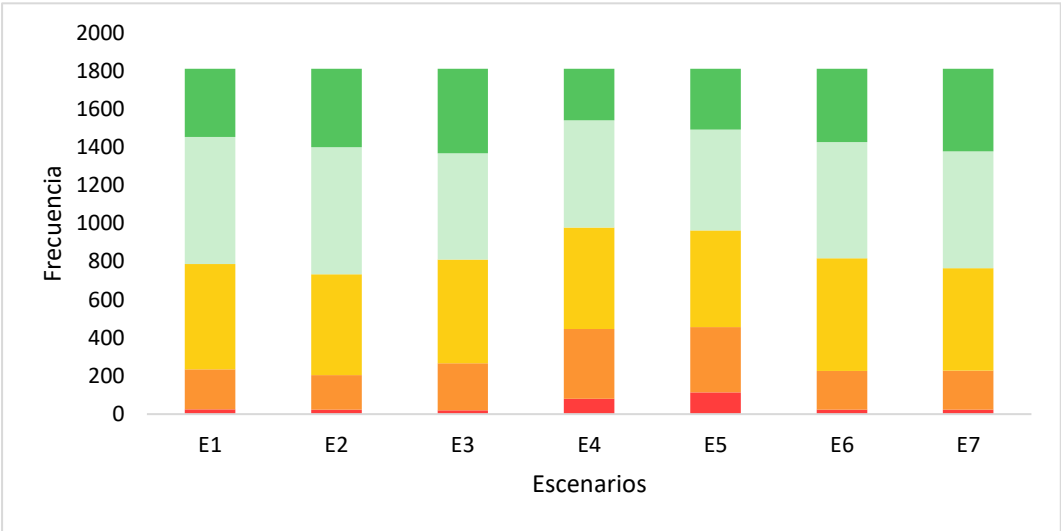


Figura 5.21. Distribución por categorías del análisis con simulación estadística

Análisis de sensibilidad basado en cambios en los pesos de las variables

En este caso, se modificaron los pesos asignados a las variables mediante un modelo cognitivo con el propósito de evaluar el impacto que tienen dichos ajustes en el proceso de toma de decisiones. Este enfoque permitió simular distintas percepciones o prioridades que podrían tener los tomadores de decisiones en contextos reales, incorporando elementos subjetivos en la ponderación de los criterios. Al variar los pesos de forma controlada, es posible analizar la sensibilidad del modelo ante cambios en la importancia relativa de cada variable, lo cual contribuye a identificar cuáles factores influyen más significativamente en los resultados y cómo podrían variar las decisiones bajo diferentes escenarios o perspectivas. Esta estrategia no solo fortalece

la robustez del modelo, sino que también mejora su aplicabilidad a contextos dinámicos o inciertos donde las preferencias pueden evolucionar con el tiempo.

Para evaluar el impacto de los cambios en los parámetros del modelo matemático multicriterio, en particular la variación en los pesos de las variables, se utilizó la herramienta Mental Modeler (Gray & Cox, 2013), una plataforma basada en diagramas causales que facilita el análisis de problemas complejos. Esta herramienta permite representar gráficamente la relación entre variables y simular distintos escenarios para evaluar posibles resultados bajo diversas condiciones y supuestos.

El uso de diagramas causales, a través de la Teoría de Gráficas ó Teoría de grafos, presenta varias ventajas. Una de ellas es que facilita una visualización estructurada del sistema; permitiendo comprender mejor las relaciones entre las variables e identificar aquellos nodos con un alto grado de influencia. Con las medidas de centralidad, es posible detectar las variables críticas con mayor impacto en el modelo, lo que ayuda a enfocar el análisis en los factores más relevantes. A continuación, se describen las medidas de centralidad utilizadas.

- Indegree: Representa la cantidad de influencias recibidas por un nodo de una gráfica, considerando la intensidad de las conexiones. Un valor alto sugiere que una variable está fuertemente condicionada por otros.
- Outdegree: Indica la cantidad de influencias que un nodo ejerce sobre otros, reflejando su capacidad de propagación dentro del sistema.
- Centralidad de grado: Mide el número total de conexiones directas de un nodo, sumando tanto influencias recibidas como emitidas. Una mayor centralidad indica un rol clave en la estructura del modelo.

Otro beneficio importante es la posibilidad de evaluar la propagación de efectos, es decir, analizar cómo un cambio en una variable afecta al resto del sistema, lo que resulta especialmente útil. Asimismo, la estructura de relaciones obtenida a través del grafo puede compararse con los pesos asignados en el modelo multicriterio, proporcionando un mecanismo adicional de validación. Finalmente, esta herramienta permite la exploración de escenarios alternativos; ya que, al modificar las conexiones o los pesos en el grafo, se pueden simular distintos escenarios hipotéticos y evaluar su impacto en la clasificación final de las alternativas.

Este proceso consistió en construir una matriz de correlación ponderada, integrando los resultados de la matriz de correlación de las variables con los valores de los pesos obtenidos en el Indicador de Accesibilidad a Servicios. Esta matriz no solo captura la interdependencia entre las variables, sino que también refleja la magnitud de la influencia de cada variable dentro del modelo. La Tabla 5.8 muestra los valores de la matriz de correlación ponderada y la Figura 5.22 muestra el grafo que representa la relación entre las variables.

Para analizar la estructura de relaciones, se representó esta matriz mediante un grafo causal, donde los nodos corresponden a las variables y los arcos reflejan sus relaciones ponderadas. Este enfoque permitió visualizar la influencia relativa de cada variable dentro de la red, identificando patrones de conexión que podrían no ser evidentes en un análisis tabular.

Para cuantificar la importancia estructural de cada variable dentro de la red, se utilizaron medidas de centralidad, ampliamente empleadas en el análisis Gráficas. La Tabla 5.9 muestra los valores obtenidos en la centralidad del grafo. Este resultado muestra que la variable PGRAPROES emerge como el nodo con mayor relevancia en términos de conectividad. Este resultado es consistente con los pesos obtenidos en el modelo multicriterio, validando su papel central en el análisis.

Tabla 5.8. Matriz de correlación ponderada

	SBASC	OVHAC	PGRAPROES	PVPH_SNBIE	PVPH_SINTIC	NUM_ESC	DIST_PR_PIE
SBASC		0.100772725	0.090489782	0.021374836	0.03055933	0.016347207	0.012921555
OVHAC	0.11234302		0.09573692	0.027339879	0.038967763	0.012154396	0.008027313
PGRAPROES	0.178028707	0.168953369		0.04259724	0.061298522	0.025018809	0.018875957
PVPH_SNBIE	0.026211126	0.030072975	0.02655057		0.092645211	-0.021041541	-0.010532787
PVPH_SINTIC	0.048418561	0.055382229	0.049365956	0.119703854		-0.030167311	-0.017908458
NUM_ESC	0.01873344	0.01249408	0.014573039	-0.019663853	-0.021819387		0.102030144
DIST_PR_PIE	0.014773728	0.008232703	0.010969676	-0.009820548	-0.012923063	0.101795793	

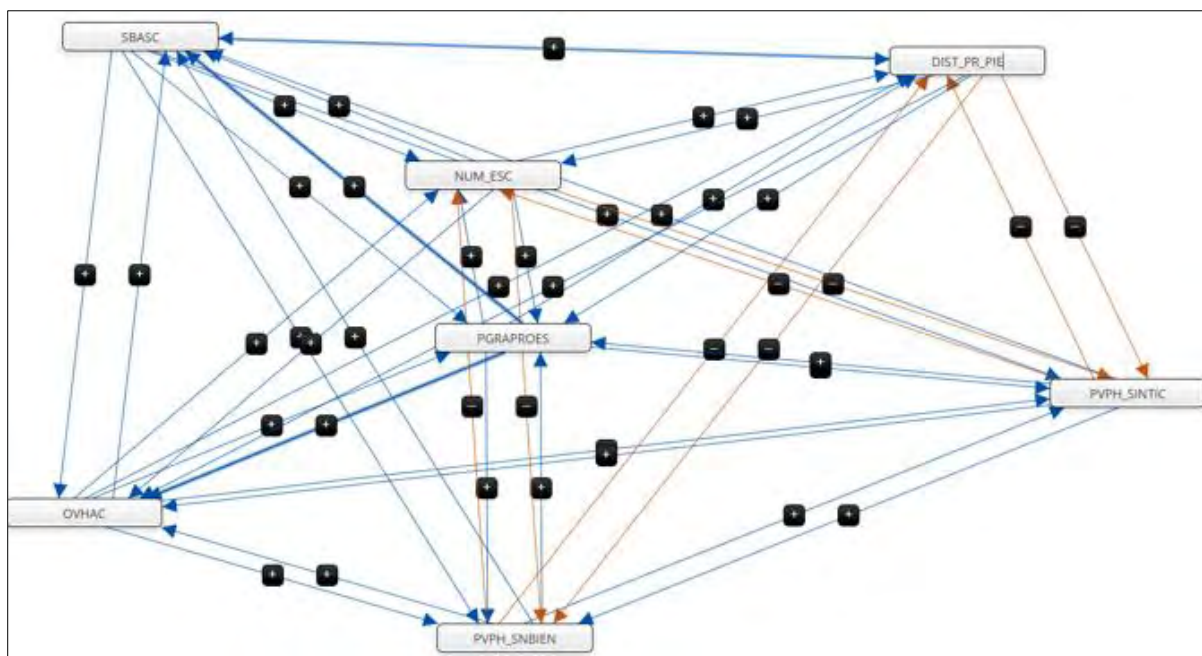


Figura 5.22. Grafo con la relación de las variables

Tabla 5.9. Medidas de centralidad del grafo

Total Components	Component	Indegree	Outdegree	Centrality
7	SBASC	1.398508582	1.2724654349999995	2.670974017
Total Connections	OVHAC	1.3759080809999997	1.2945692909999997	2.6704773719999997
49	PGRAPROES	1.2876859429999998	1.494772604	2.782458547
Density	PVPH_SNBIE	1.24050021	1.2070542100000001	2.4475544200000003
1.166666667	PVPH_SINTIC	1.258213276	1.320946369	2.579159645
Connections per Component	NUM_ESC	1.206525057	1.189313943	2.3958389999999996
7	DIST_PR_PIE	1.170296214	1.158515511	2.328811725

La simulación de escenarios implicó introducir variaciones en los pesos de las variables para observar cómo impactan en algunas de las variables de modelo. A continuación, se describen algunos resultados de la simulación de escenarios ante variaciones de los pesos de las variables.

Escenario 1: Se modificó el valor de la variable NUM_ESC a -0.32, correspondiente al (número de escuelas). Este cambio generó un impacto positivo en las variables PVPH_SNBIEN (porcentaje de viviendas sin acceso a bienes básicos) y PVPH_SINTIC. En la Figura 5.23., las barras de color rojo indican que la variable PVPH_SNBIEN (porcentaje de viviendas sin acceso a tecnologías) incrementó a un valor de 0.01 y PVPH_SINTIC subió a un valor de 0.01 por encima del eje x. Este resultado indica una relación inversa entre el número de escuelas y las condiciones de la población con el acceso a los bienes y la tecnología. Es decir, si hay menos escuelas, aumenta el número de viviendas sin acceso a las tecnologías y bienes.

Escenario 2: Se modificó el valor de las variables PVPH_SNBIEN en -0.3, (correspondiente a las viviendas sin acceso a bienes). Esta modificación generó un impacto negativo y positivo en las variables (viviendas sin acceso a bienes, SBASC (porcentaje de personas sin educación básica), OVHAC (porcentaje de personas que viven en condiciones de hacinamiento) y PVPH_SINTIC (porcentaje de viviendas sin acceso a tecnologías). En la Figura 5.24 las barras indican un incremento o disminución de los valores de las variables sobre el eje x. En la figura se observa una disminución en los valores de las variables SBASC (población sin educación básica), OVHAC (población con hacinamiento) y PGRAPROES (grado de escolaridad) a -0.01 respectivamente. Por otro lado, hubo un incremento en el valor de la variable NUM_ESC (número de escuelas) de 0.01. Este resultado indica si hay menos viviendas sin acceso a bienes, también hay menos viviendas con acceso a las tecnologías, menos población sin acceso a la educación básica y menos población sin hacinamiento. Asimismo, si hay menos viviendas sin acceso a bienes, hay mayor número de escuelas.

Escenario 3: Se modificó el valor de la variable NUM_ESC en 0.2, (correspondiente a al número de escuelas). Este cambio generó un impacto negativo en la variable DIST_PR_PIE (distancia promedio de viaje). En la Figura 5.25 la barra indica una disminución del valor de la variable sobre el eje x. En la figura se observa una disminución de DIST_PR_PIE a -0.01. Este resultado indica si hay más escuelas la distancia del tiempo de viaje disminuye.

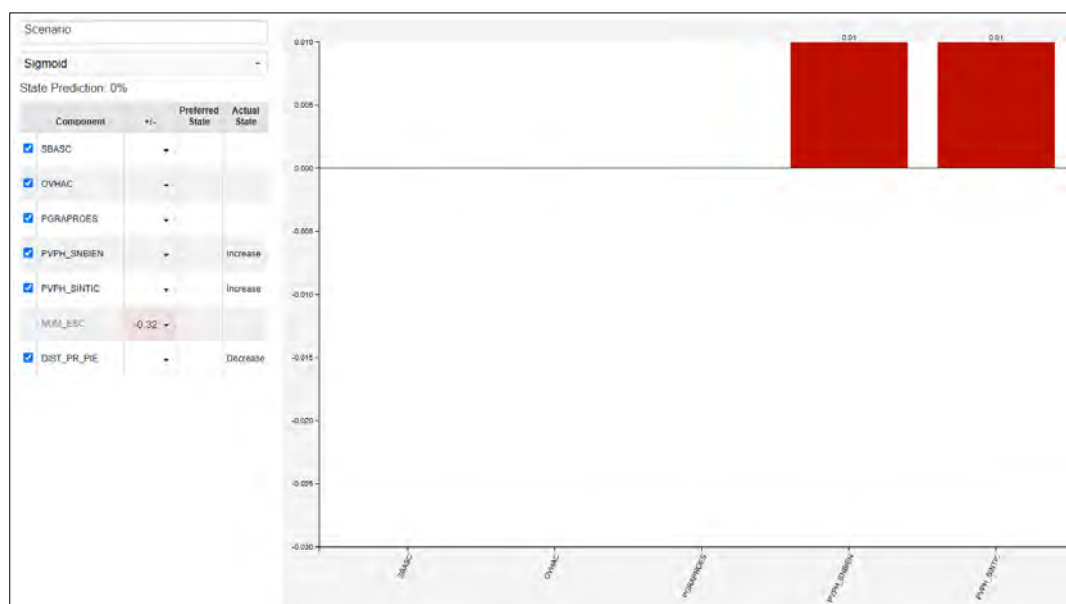


Figura 5.23. Simulación del escenario 1



Figura 5.24. Simulación del escenario 2

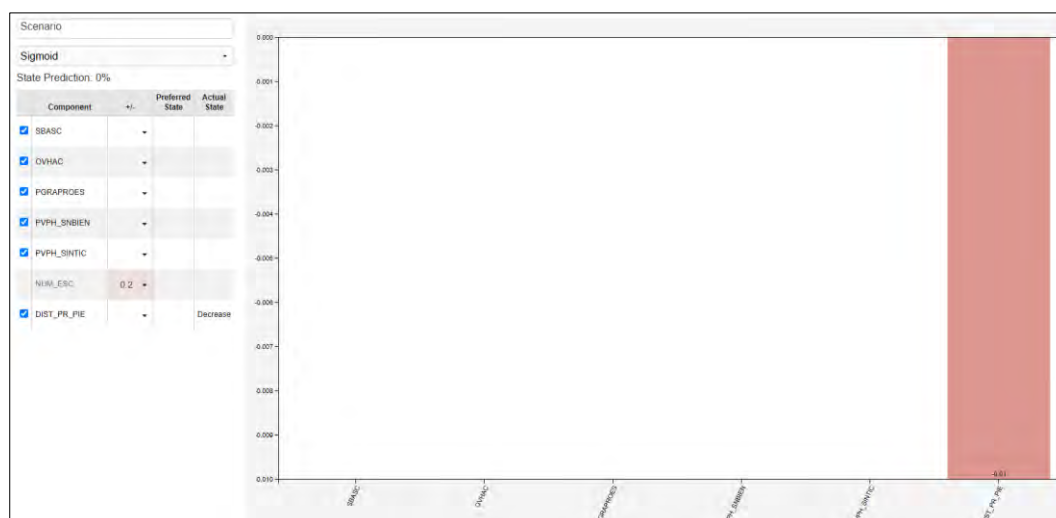


Figura 5.25. Simulación del escenario 3

5.1.3. Análisis de ablación

El análisis de ablación consiste en eliminar sistemáticamente una o más variables del modelo para evaluar su impacto en el desempeño del mismo. Este enfoque permite identificar la relevancia de cada componente del modelo al observar cómo cambian los resultados al prescindir de ciertas variables o dimensiones. En contextos de la Toma de Decisiones Multicriterio, como son indicadores compuestos, el análisis de ablación resulta especialmente útil para validar la robustez del modelo, detectar posibles redundancias y comprender mejor el papel o importancia que tiene cada dimensión en la construcción del indicador final (Jiang et al., 2025), (Park et al., 2025).

En esta investigación, el estudio de ablación consistió en eliminar cada una de las tres dimensiones o criterios de educación, vivienda y espacial que conforman modelo matemático multicriterio propuesto, definido como Indicador de Accesibilidad a Servicios. Esto con el objetivo de analizar su influencia individual sobre el valor del indicador. Mediante el coeficiente de correlación de Spearman y a través de esta técnica, se evaluó cómo la exclusión de cada dimensión afecta la correlación del IAS con respecto de cada uno de los modelos de referencia: el Índice de Marginación, el Índice de Oportunidades y el Índice de Accesibilidad basada en distancia. Este enfoque permitió identificar qué dimensiones aportan mayor coherencia estructural al modelo y cuáles son indispensables para reflejar de manera integral las condiciones de acceso en el territorio analizado.

En el análisis de ablación realizado, la exclusión del criterio espacial se identificó como particularmente influyente en la alineación del modelo propuesto con el Índice de Marginación. Al eliminar esta dimensión del modelo, la correlación con dicho índice aumentó de 0.5829 (modelo completo) a 0.6113. Esto indica que los componentes espaciales, como la distancia promedio a pie o la cantidad de escuelas, introducen información que no está plenamente capturada por el índice de marginación. No obstante, esta mejora en la alineación con el índice de marginación ocurre a costa de una fuerte pérdida de coherencia con los otros dos indicadores de referencia: el

indicador de oportunidades acumuladas (-0.0921) y el indicador basado en la distancia (-0.0925). Esta caída abrupta evidencia que la dimensión espacial es indispensable para reflejar con precisión la accesibilidad territorial real y las oportunidades geográficas disponibles para la población.

Por otro lado, la exclusión del criterio de vivienda resultó ser la más relevante en términos de coherencia estructural general del modelo con los tres indicadores externos. Al eliminar esta dimensión, la correlación con el índice de marginación disminuyó a 0.4854, mientras que con el indicador de oportunidades se redujo a -0.5832 y con el indicador basado en la accesibilidad por distancia a -0.6762; siendo esta última la correlación negativa más fuerte observada en el estudio. Estos resultados indican que las variables asociadas a las condiciones de vivienda están fuertemente alineadas con situaciones de marginación y desigualdad en el acceso a servicios, y que su exclusión debilita significativamente la capacidad explicativa modelos propuestos frente a estos fenómenos.

En cuanto a la exclusión del criterio de educación, su impacto, aunque presente, es comparativamente menor. La eliminación de esta dimensión redujo la correlación con el índice de marginación a 0.4393, y con los indicadores de oportunidades y basado en la distancia a -0.4896 y -0.6097, respectivamente. Si bien estos valores siguen reflejando una pérdida de alineación, el efecto es menos drástico que el observado al omitir la dimensión de vivienda. Esto sugiere que, aunque la educación representa un factor relevante en términos de accesibilidad, su peso estructural en el modelo es más limitado frente al impacto de las dimensiones espacial y de vivienda. La Tabla 5.10 muestra los resultados.

En síntesis, los resultados del estudio de ablación permiten concluir que el criterio de vivienda es el más determinante para mantener la coherencia del modelo matemático multicriterio propuesto con respecto de los indicadores de referencia utilizados. El criterio espacial es clave para capturar la accesibilidad efectiva desde una perspectiva territorial; mientras que la dimensión educativa, aunque importante, presenta un impacto más moderado. Este análisis respalda la inclusión de las tres dimensiones en el modelo final, ya que en conjunto permiten garantizar un equilibrio entre estructura social, bienes y servicios básicos en la vivienda y condiciones geográficas. Además, refuerza la utilidad del modelo matemático multicriterio como herramienta para el diagnóstico territorial, al integrar componentes que no son plenamente reflejados por los modelos tradicionales, lo que resulta clave para la formulación de políticas públicas orientadas a la equidad en el acceso a servicios.

Tabla 5.10 Resultados del estudio de ablación

	Modelo matemático multicriterio propuesto (Indicador de Accesibilidad a Servicios)		
	Sin educación	Sin vivienda	Sin espacial
Índice de Marginación	$\rho = 0.4393$ $p - valor = 2.508e - 86$	$\rho = 0.4854$ $p - valor = 1.143e - 107$	$\rho = 0.6113$ $p - valor = 4.518e - 186$
Indicador de oportunidades acumuladas	$\rho = -0.4896$ $p - valor = 8.176e - 110$	$\rho = -0.5832$ $p - valor = 1.675e - 165$	$\rho = -0.0921$ $p - valor = 8.604e - 05$
Indicador basado en la distancia	$\rho = -0.6097$ $p - valor = 8.689e - 185$	$\rho = -0.6762$ $p - valor = 2.438e - 242$	$\rho = -0.0925$ $p - valor = 8.012e - 05$

5.2. Interpretación del modelo

La interpretación del modelo multicriterio, definido como Indicador de Accesibilidad a Servicios, implica analizar los valores de salida evaluados y utilizar técnicas para su representación cualitativa y cuantitativa (CONEVAL, 2014). Este proceso facilita una comprensión más profunda del modelo matemático desarrollado; permitiendo identificar hallazgos relevantes que contribuyen a una toma de decisiones informadas. A través de este análisis, se pueden descubrir patrones, tendencias y relaciones significativas entre las múltiples variables involucradas; Así como establecer comparaciones sobre el desempeño de las distintas unidades de análisis (alternativas) a partir de los cálculos realizados (Schuschny & Soto, 2009).

Esta etapa implica desglosar y comprender cómo cada alternativa se posiciona frente a las variables establecidas; permitiendo identificar no solo las zonas con mejor acceso, sino también aquellas que presentan mayores dificultades. Esta evaluación proporciona una base sólida para la toma de decisiones, al ofrecer una visión clara de las disparidades en el acceso a los servicios y facilitar la identificación de áreas prioritarias para su intervención.

Clasificación de los niveles de accesibilidad

El primer paso consistió en interpretar los valores de salida del modelo, a partir de la evaluación de las alternativas en el modelo multicriterio. Los valores fueron clasificados en categorías que representan los niveles de accesibilidad por localidad hacia las escuelas primarias, alineados con los niveles utilizados en indicadores sociales presentados en fuentes oficiales. Estos indicadores utilizan una escala de cinco niveles: Muy bajo, Bajo, Medio, Alto y Muy alto. La clasificación de los niveles de accesibilidad se determinó mediante la evaluación de las alternativas en el modelo y la aplicación posterior del método de optimización de Jenks. Como se mencionó anteriormente, este método agrupa los datos de manera óptima, minimizando la varianza dentro de cada clase y maximizando la diferencia entre clases. Dicho proceso se realizó durante el desarrollo del modelo. A continuación, se presentan las reglas establecidas para determinar los niveles de accesibilidad de las localidades; las cuales permiten medir cuantitativa y cualitativamente el nivel de accesibilidad que cada localidad tiene hacia las escuelas de nivel primaria, véase Tabla 5.11. Esta combinación de categorización,

clasificación y asignación de colores mejora la interpretación visual y la toma de decisiones.

Por otro lado, Figura 5.26 ilustra el histograma con la distribución de las localidades según su nivel de accesibilidad, clasificado en cinco categorías: muy alta, alta, media, baja y muy baja; cada una representada por un color distintivo para facilitar su interpretación. En el eje X se encuentran los intervalos de clase correspondientes a los valores del Indicador de Accesibilidad a Servicios; mientras que el eje Y refleja la frecuencia de localidades en cada rango. La mayor concentración de localidades se sitúa en los rangos negativos y cercanos a cero; correspondientes a las categorías de muy alta y alta accesibilidad, lo que sugiere que una parte significativa de las localidades cuenta con buen acceso a los servicios esenciales. A medida que los valores se desplazan hacia la derecha del gráfico, se observa una disminución en la frecuencia, representando niveles de accesibilidad media y baja. Las localidades con muy baja accesibilidad son escasas, lo que indica que una minoría enfrenta serias limitaciones en el acceso a los servicios.

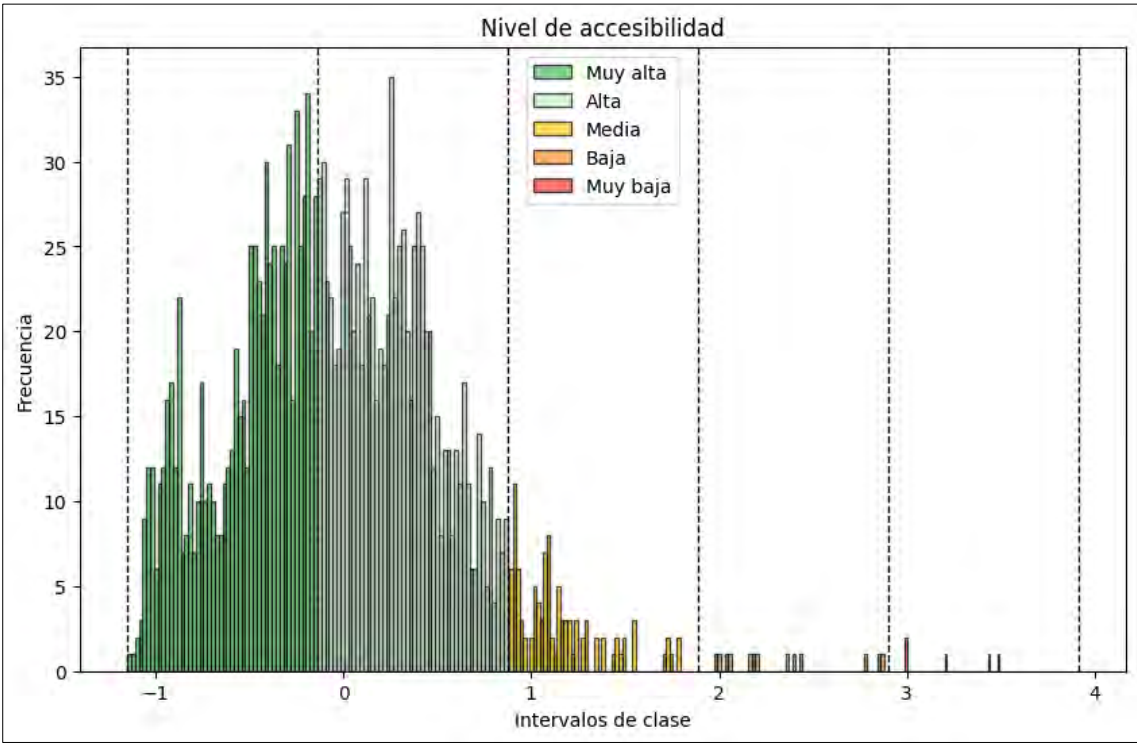


Figura 5.26. Histograma de los intervalos de clase clasificador mediante Jenks

Tabla 5.11. Reglas de la escala numérica y cualitativa del nivel de accesibilidad

Color	Valor de la escala	Nivel de accesibilidad
	1	Accesibilidad Muy Baja
	2	Accesibilidad Baja
	3	Accesibilidad Media
	4	Accesibilidad Alta
	5	Accesibilidad Muy Alta

Todo este proceso se refleja en un árbol de decisión; el cual proporciona una representación clara, estructurada y de fácil interpretación de la accesibilidad hacia las escuelas primarias; facilitando su comprensión y análisis. A continuación, se describen los elementos del árbol de decisión.

- o **Nodo raíz:** Es el nodo principal del árbol, el cual representa la variable de decisión sobre la que se basa la clasificación, definida como Nivel de accesibilidad a servicios. A partir de este nodo se establecen los rangos numéricos que determinan el nivel de accesibilidad de manera cuantitativa y cualitativa.
- o **Valor de Decisión:** El árbol de decisión incorpora una estructura que permite calcular una puntuación global del Indicador de Accesibilidad a Servicios. En este proceso, cada localidad ingresa al sistema y es evaluada de manera integral con base en múltiples variables representados por diversas variables. Cada variable contribuye al desempeño total en función de su valor y peso específico dentro del modelo. A partir de la puntuación obtenida, la localidad se clasifica en uno de los cinco niveles de accesibilidad, de acuerdo con las reglas establecidas; lo que permite interpretar de manera clara su nivel de accesibilidad a las escuelas primarias, véase Tabla 5.12.

El árbol de decisión actúa como una herramienta de evaluación multicriterio que facilita la clasificación y análisis de la accesibilidad a las escuelas de nivel, ofreciendo una estructura clara y jerárquica para evaluar la accesibilidad primaria y permitiendo una visión más completa y estratégica para la toma de decisiones. La Figura 5.27 muestra el esquema del árbol de decisión.

Tabla 5.12. Valor de decisión

Color	Intervalos de valores	Categoría	Descripción del Nivel de Accesibilidad
	[1.7315, 4]	1	La accesibilidad a las escuelas primarias en la zona de interés es Muy Baja
	[0.5611, 1.7135)	2	La accesibilidad a las escuelas primarias en la zona de interés es Baja
	[-0.0017, 0.5611)	3	La accesibilidad a las escuelas primarias en la zona de interés es Media
	[-0.5282, -0.0017)	4	La accesibilidad a las escuelas primarias en la zona de interés es Alta
	[-1.2, -0.5282)	5	La accesibilidad a las escuelas primarias en la zona de interés es Muy Alta

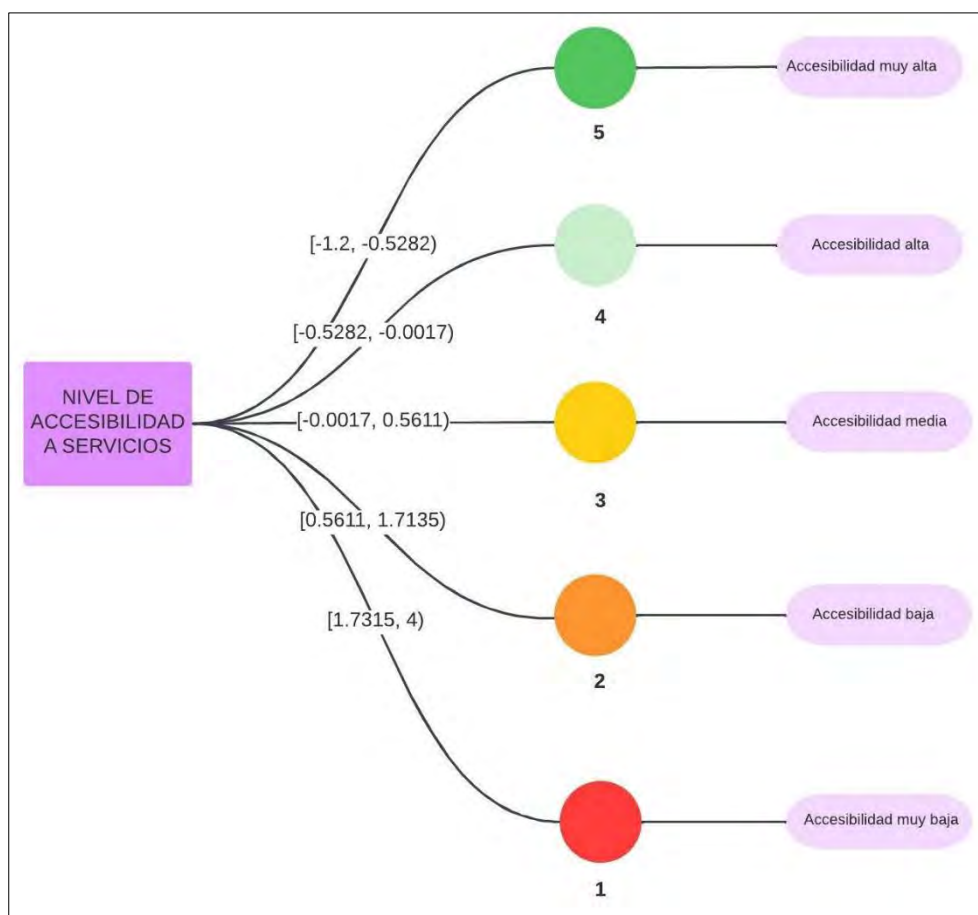


Figura 5.27. Árbol de decisión del indicador de accesibilidad a servicios

5.2.1. Conjunto de datos evaluado

A partir de lo anterior, se obtuvo un conjunto de datos en el que cada registro representa una localidad evaluada con base en diversas variables, asignándole un nivel de accesibilidad según el indicador calculado, véase Tabla 5.13. Este conjunto de datos incluye información geoespacial de las localidades, los valores de las variables consideradas, el valor obtenido del indicador y su correspondiente nivel de accesibilidad. Cada registro contiene información clave que influye en la evaluación de la accesibilidad a servicios, permitiendo un análisis integral de cada área. Los campos incluidos en la tabla se describen en la Tabla 5.15.

Por su parte, La Tabla 5.14 presenta la distribución de localidades según los cinco niveles de accesibilidad, brindando una visión general del acceso a servicios en el territorio. El análisis de estos datos indica que la mayoría de las localidades se agrupan en los niveles de accesibilidad media y alta, con un total de 1,811 localidades, lo que sugiere un panorama favorable en términos de acceso a servicios. No obstante, se identifican 270 localidades en los niveles de accesibilidad baja y muy baja, lo que resalta la necesidad de atención prioritaria para mejorar sus condiciones.

Tabla 5.13. Conjunto de datos del modelo de accesibilidad

FID	NOMDT	NOMUT	SBASC	OVHAC	PGRAPR OES	PVPH_SN BIEN	PVPH_SI NTIC	NUM_ESC	DIST_PR_PIE	IND ACC
1	AZCAPOTZALCO	AGUILERA	16.378	24.966	8.786	0.000	0.000	1.000	207.000	4
2	AZCAPOTZALCO	ALDANA	16.537	29.071	9.220	0.000	0.265	10.000	10000.000	3
3	AZCAPOTZALCO	ANGEL ZIMBRON	10.531	14.337	8.670	0.000	0.000	0.333	195.667	5
4	AZCAPOTZALCO	ARENAL	14.658	22.528	9.292	0.000	0.222	1.000	481.000	4
5	AZCAPOTZALCO	CENTRO DE AZCAPOTZALC O	12.231	18.308	9.283	1.137	1.487	0.250	306.250	2
6	AZCAPOTZALCO	CLAVERIA	8.199	8.925	7.285	0.000	0.000	0.143	463.143	5
7	AZCAPOTZALCO	COLONGO	23.040	28.622	10.358	0.000	0.000	1.000	278.000	3
8	AZCAPOTZALCO	COSMOPOLITA	17.175	23.671	8.760	0.000	0.000	1.000	222.000	4
9	AZCAPOTZALCO	COSMOPOLITA (AMPL)	9.916	14.814	11.819	0.000	0.000	1.000	27.000	4
10	AZCAPOTZALCO	CRUZ ROJA TEPANTONGO (U HAB)	17.038	21.795	9.112	0.000	0.000	10.000	10000.000	3
11	AZCAPOTZALCO	CUITLÁHUAC 1 Y 2 (U HAB)	9.604	6.867	7.563	0.000	0.000	0.500	912.500	5
12	AZCAPOTZALCO	CUITLÁHUAC 3 Y 4 (U HAB)	9.604	6.867	7.818	0.000	0.000	10.000	10000.000	4
13	AZCAPOTZALCO	DEL GAS	12.736	19.608	8.167	0.000	0.000	0.333	219.000	5
14	AZCAPOTZALCO	DEL GAS (AMPL)	10.979	17.663	7.114	0.000	0.000	0.333	220.667	5
15	AZCAPOTZALCO	DEL MAESTRO	17.091	21.586	7.953	0.000	0.000	10.000	10000.000	3
16	AZCAPOTZALCO	DEL RECREO	9.010	11.565	7.801	0.000	0.000	0.333	268.333	5
17	AZCAPOTZALCO	DEMET (U HAB)	11.675	16.403	7.696	0.000	0.000	10.000	10000.000	4
18	AZCAPOTZALCO	ECOLOGICA NOVEDADES IMPACTO (U HAB)	16.378	24.966	8.786	0.000	0.000	1.000	207.000	4

Tabla 5.14. Reporte del número de localidades por categoría

Número de localidades	Nivel de accesibilidad	
25	1	Muy baja
245	2	Baja
588	3	Media
631	4	Alta
325	5	Muy alta

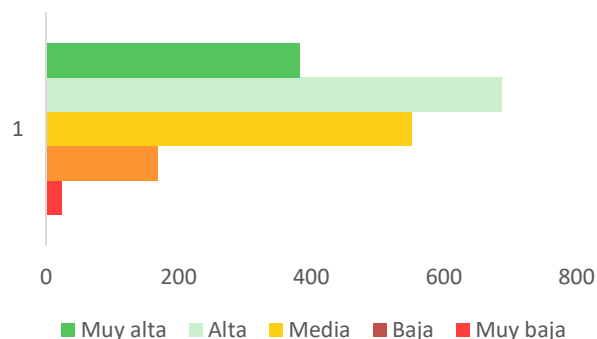


Tabla 5.15. Descripción del conjunto de datos de accesibilidad a las escuelas

VARIABLE	DESCRIPCIÓN	TIPO DE DATO
ID	Id de la localidad	-
CVEDT	Clave del municipio	Carácter
NOMDT	Nombre del municipio	Texto
CVEUT	Clave de la localidad	Carácter
NOMUT	Nombre de la localidad	Texto
X	Longitud (ubicación geográfica de la localidad)	Continúo
Y	Latitud (ubicación geográfica de la localidad)	Continúo
SBASC	Porcentaje de la población de 15 años o más sin educación básica	Continúo
OVHAC	Porcentaje de viviendas con hacinamiento	Continúo
PGRAPROES	Grado promedio de escolaridad	Continúo
PVPH_SNBIEN	Porcentaje de viviendas habitadas sin ningún bien	Continúo
PVPH_SINTIC	Porcentaje de viviendas sin tecnologías de la información y de la comunicación (TIC)	Continúo
NUM_ESC	Número de escuelas en la localidad	Entero
TIEM_PR_PIE	Tiempo promedio de viaje del centroide de la localidad hacia la ubicación geográfica de las escuelas en la modalidad a pie.	Continúo (minutos)
DIST_PR_PIE	Distancia promedio del trayecto del centroide de la localidad hacia la ubicación geográfica de las escuelas en la modalidad a pie.	Continúo (metros)
IAS_EVA	Indicador de accesibilidad evaluado DE la localidad	Continúo
NIVEL_IAS	Nivel de accesibilidad asignado a la localidad (escala de 1 a 5).	Entero

5.2.2. Representación gráfica

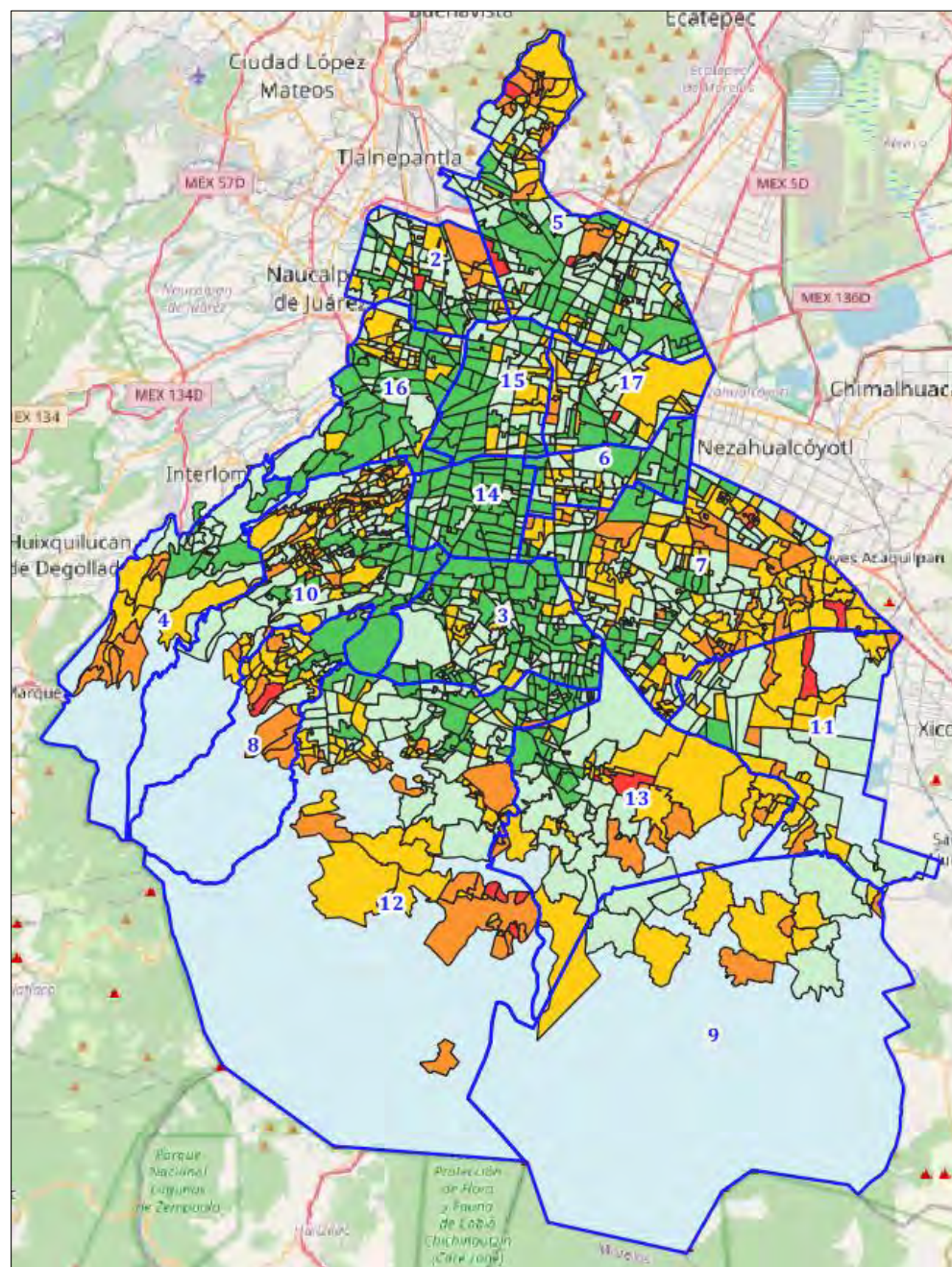
Posteriormente, utilizando el Sistema de Información Geográfica QGIS, se generó una representación gráfica del indicador de accesibilidad, para obtener una visualización clara y detallada de la distribución espacial de la distribución de los servicios y el comportamiento de la accesibilidad en función de sus niveles. Esta representación integró la información geoespacial de los municipios y localidades, la ubicación de las escuelas primarias y los niveles del indicador de accesibilidad, los cuales fueron plasmados a través de mapa coropletras para su representación cartográfica.

La Figura 5.28 muestra el mapa del indicador de accesibilidad a nivel municipal para la Ciudad de México, proporcionando una visión general de las diferencias espaciales en el acceso a los servicios educativos. En la Figura 5.29, se observa la distribución de las escuelas en relación con los niveles de accesibilidad, lo que permite identificar zonas con mayor o menor cobertura de infraestructura educativa. Finalmente, Figura 5.30, Figura 5.31, Figura 5.32, Figura 5.33, Figura 5.34, Figura 5.35, Figura 5.36, Figura 5.37, Figura 5.38, Figura 5.39, Figura 5.40, Figura 5.41, Figura 5.42, Figura 5.43, Figura 5.44, Figura 5.45 presentan un conjunto detallado de mapas con la clasificación de accesibilidad para cada municipio, facilitando el análisis comparativo entre diferentes

áreas y la identificación de regiones con mayores necesidades de mejora en términos de acceso a la educación.

El análisis de estos mapas permite identificar áreas prioritarias para la mejora de la infraestructura educativa y el acceso a servicios. En general, las zonas centrales y del norte presentan niveles altos de accesibilidad, reflejados en una mayor densidad de escuelas y mejores condiciones de conectividad. En contraste, las regiones del sur y suroeste muestran accesibilidad baja o muy baja, posiblemente debido a la geografía montañosa o una menor densidad poblacional. Además, en el oriente de la ciudad, cerca de los límites con el Estado de México, se observan áreas con accesibilidad media o baja, lo que sugiere la necesidad de fortalecer la infraestructura educativa. Estos mapas proporcionan una base sólida para el análisis territorial y la toma de decisiones en materia de planificación educativa. Al visualizar la accesibilidad a nivel espacial, se facilita la identificación de áreas críticas y la implementación de estrategias para mejorar la equidad en el acceso a la educación en la Ciudad de México.

Esta representación cartográfica no solo permite analizar la accesibilidad de manera visual e intuitiva, sino que también constituye una herramienta fundamental para la toma de decisiones en planificación territorial y políticas públicas, orientadas a mejorar la equidad en la distribución de los servicios educativos. Al visualizar la accesibilidad a nivel espacial, se facilita la identificación de áreas críticas y la implementación de estrategias para garantizar un acceso equitativo a la educación en la Ciudad de México.



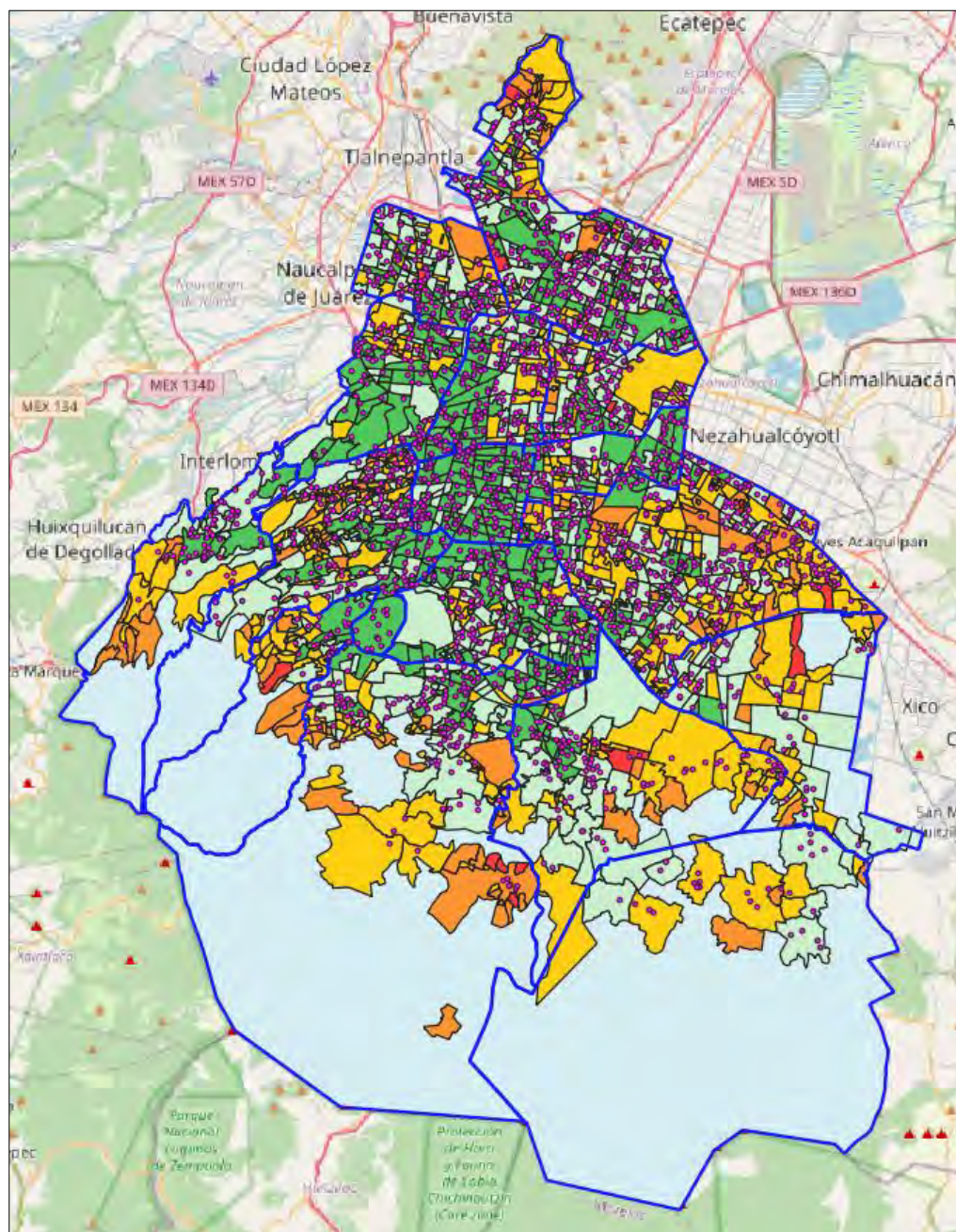
**(Clave)
Municipios**

- 2 Azcapotzalco
- 3 Coyoacán
- 4 Cuajimalpa de Morelos
- 5 Gustavo A. Madero
- 6 Iztacalco
- 7 Iztapalapa
- 8 La Magdalena Contreras
- 9 Milpa Alta
- 10 Álvaro Obregón
- 11 Tláhuac
- 12 Tlalpan
- 13 Xochimilco
- 14 Benito Juárez
- 15 Cuauhtémoc
- 16 Miguel Hidalgo
- 17 Venustiano Carranza

Nivel de accesibilidad por localidad

Color	Intervalos	Categoría	Descripción
	[1.7315, 4]	1	Muy Baja
	[0.5611, 1.7135)	2	Baja
	[-0.0017, 0.5611)	3	Media
	[-0.5282, -0.0017)	4	Alta
	[-1.2, -0.5282)	5	Muy Alta

Figura 5.28. Nivel de accesibilidad en la Ciudad de México por localidad



Nivel de accesibilidad por localidad

Color	Intervalos	Categoría	Descripción
	[1.7315, 4]	1	Muy Baja
	[0.5611, 1.7135)	2	Baja
	[-0.0017, 0.5611)	3	Media
	[-0.5282, -0.0017)	4	Alta
	[-1.2, -0.5282)	5	Muy Alta

Escuelas de nivel primaria



Figura 5.29. Nivel de accesibilidad por localidad con las escuelas

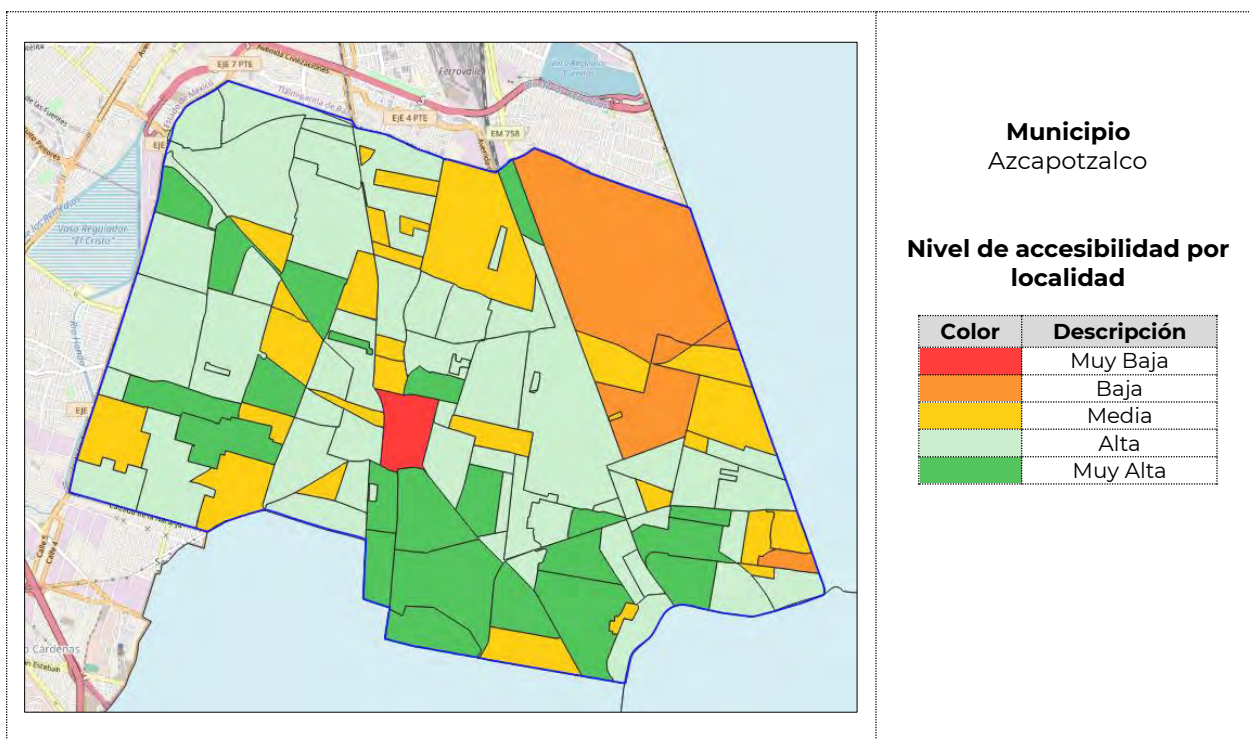


Figura 5.30. Indicador de accesibilidad: municipio de Azcapotzalco

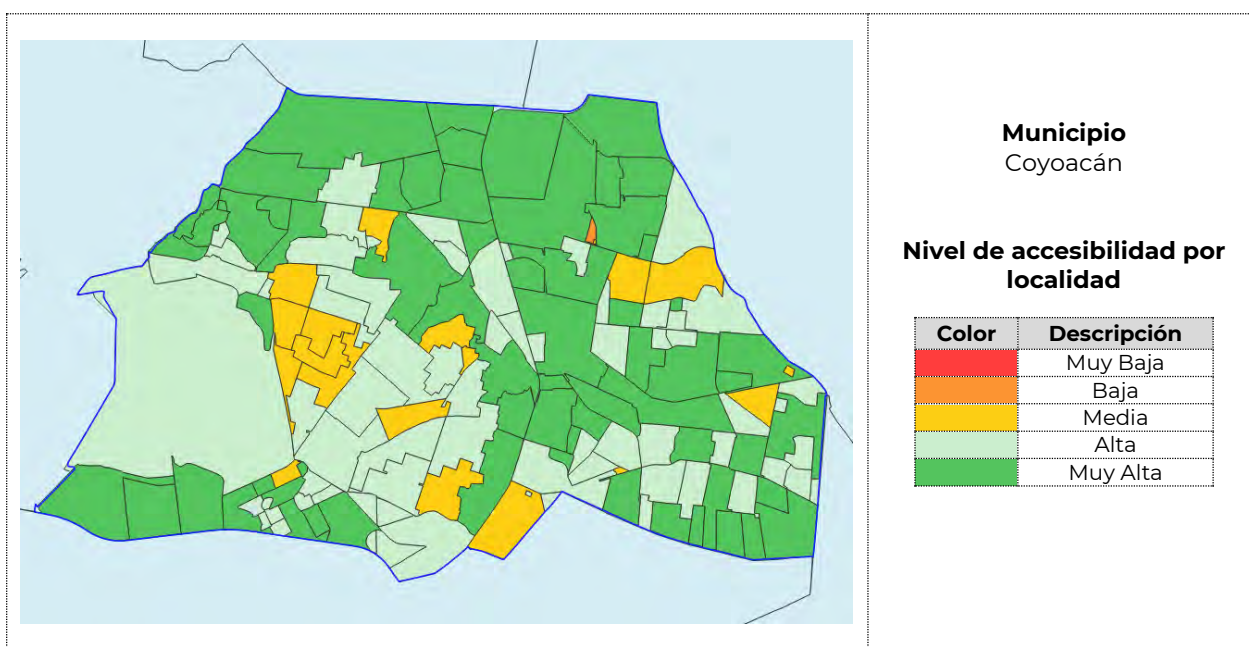


Figura 5.31. Indicador de accesibilidad: municipio de Coyoacán

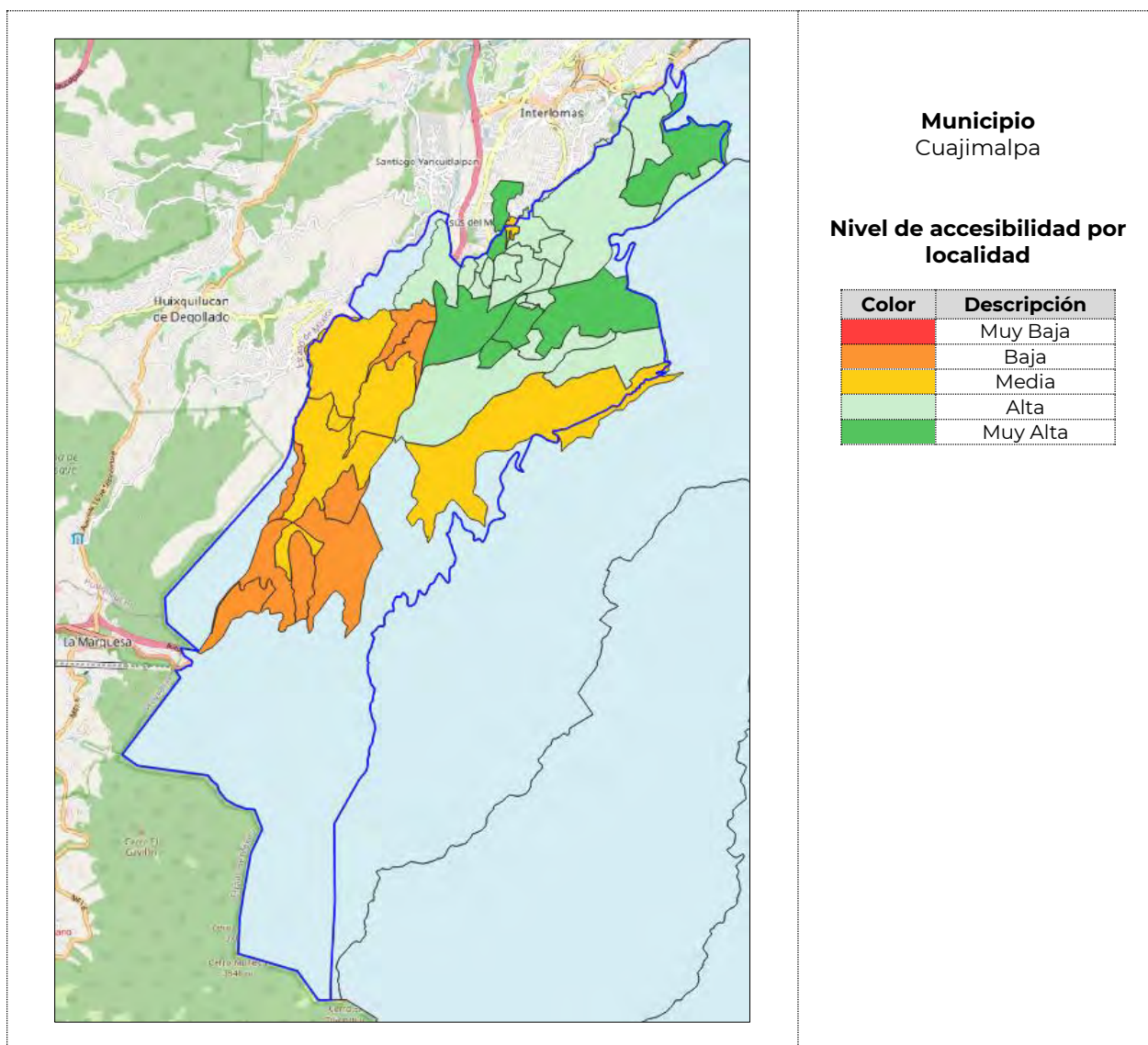
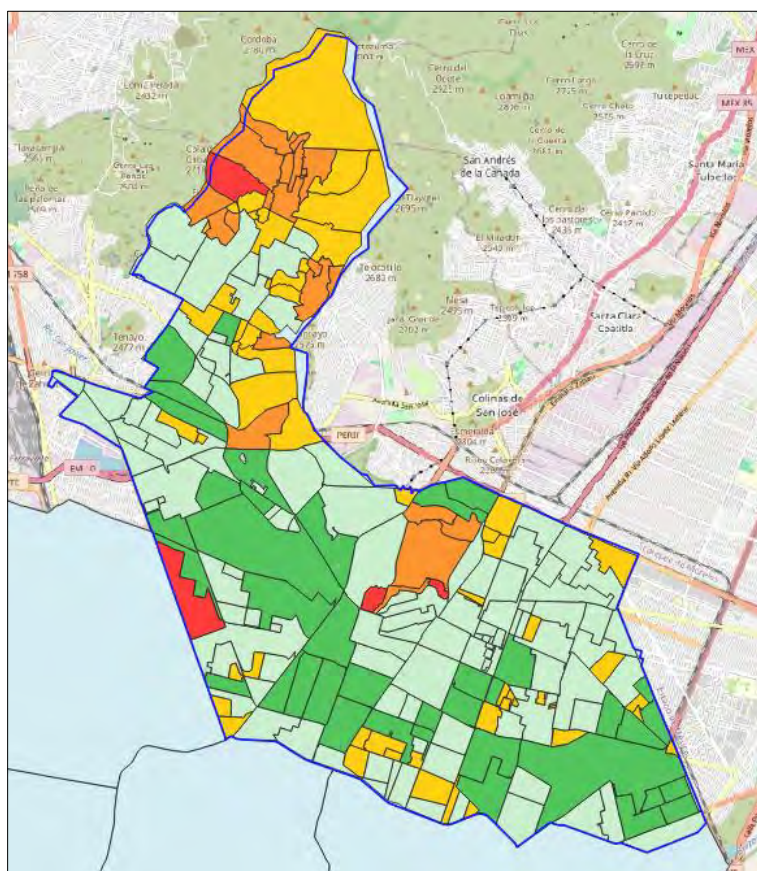


Figura 5.32. Indicador de accesibilidad: del municipio de Cuajimalpa

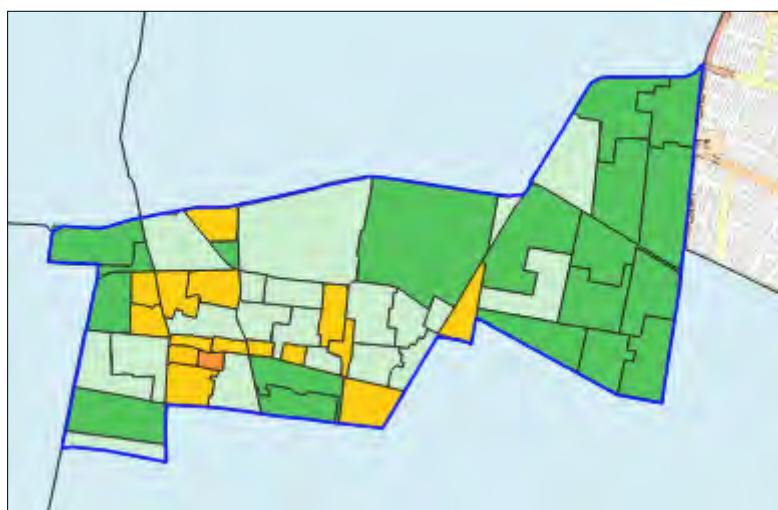


Municipio
Gustavo A. Madero

Nivel de accesibilidad por localidad

Color	Descripción
Red	Muy Baja
Orange	Baja
Yellow	Media
Light Green	Alta
Dark Green	Muy Alta

Figura 5.33. Indicador de accesibilidad: municipio de Gustavo A. Madero



Municipio
Iztacalco

Nivel de accesibilidad por localidad

Color	Descripción
Red	Muy Baja
Orange	Baja
Yellow	Media
Light Green	Alta
Dark Green	Muy Alta

Figura 5.34. Indicador de accesibilidad: municipio de Iztacalco

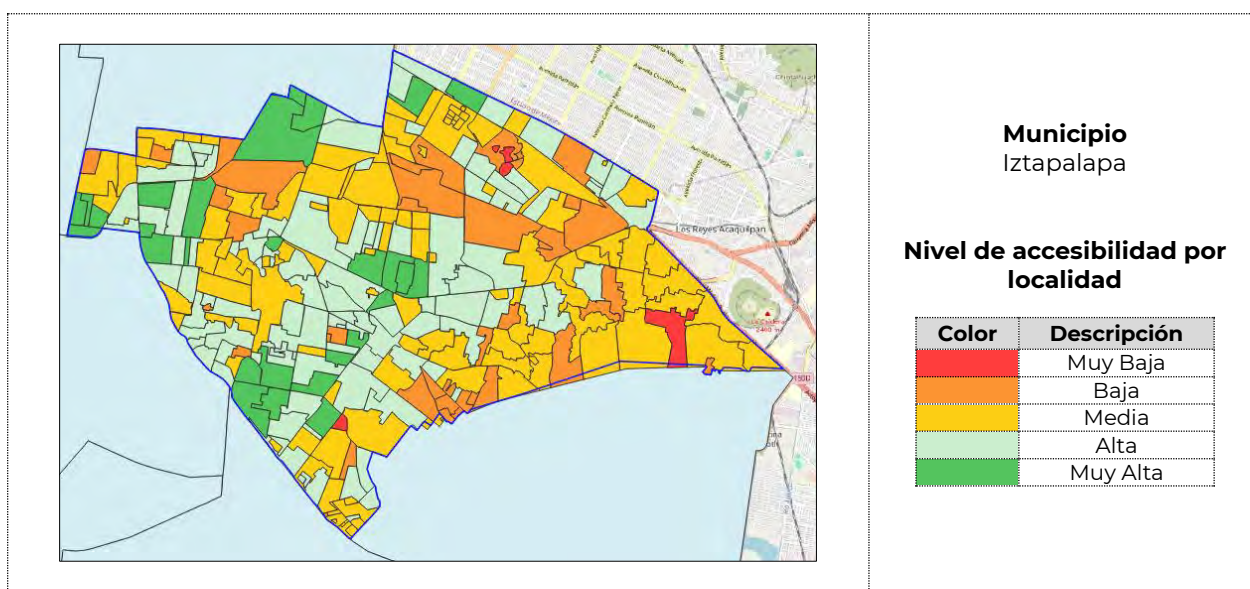


Figura 5.35. Indicador de accesibilidad: municipio de Iztapalapa

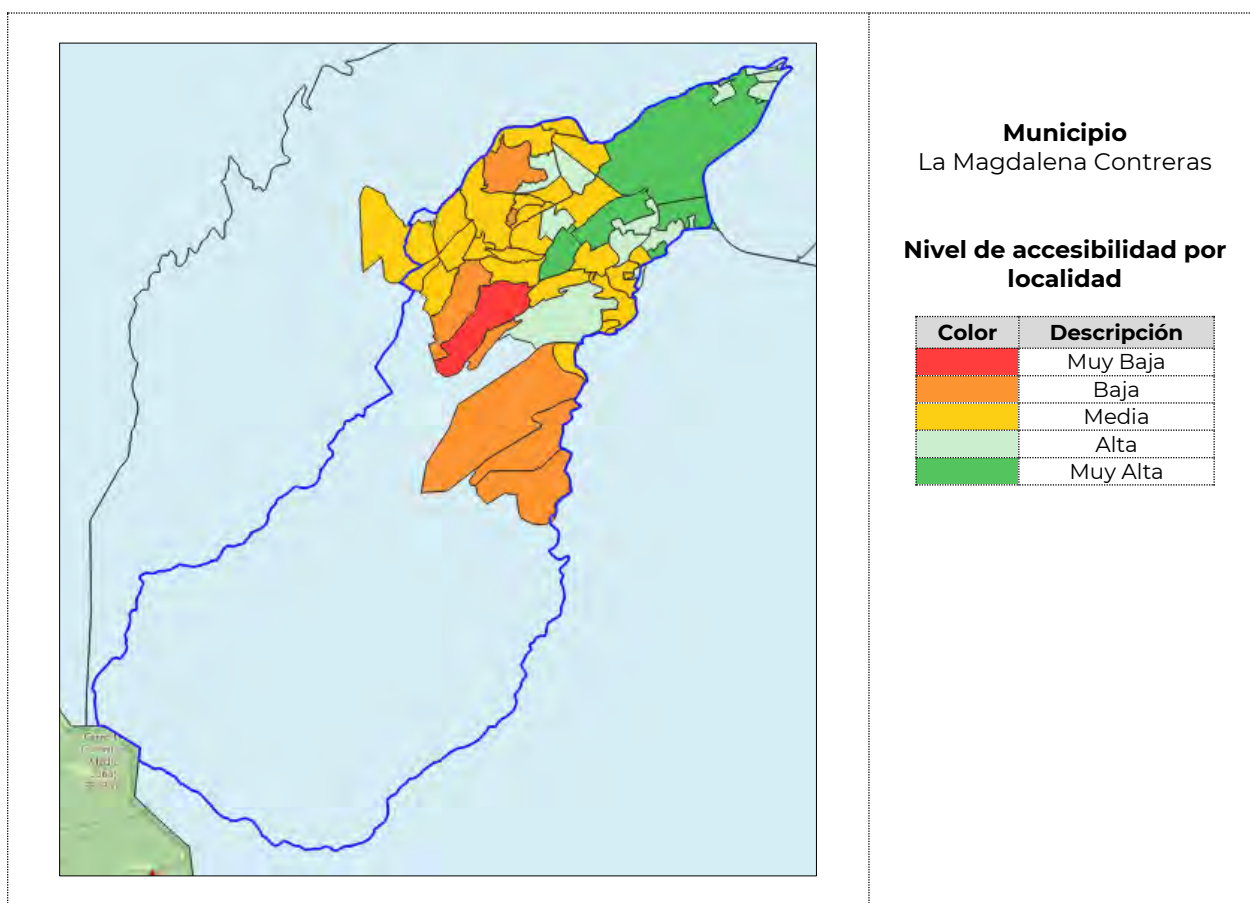


Figura 5.36. Indicador de accesibilidad: municipio de La Magdalena Contreras

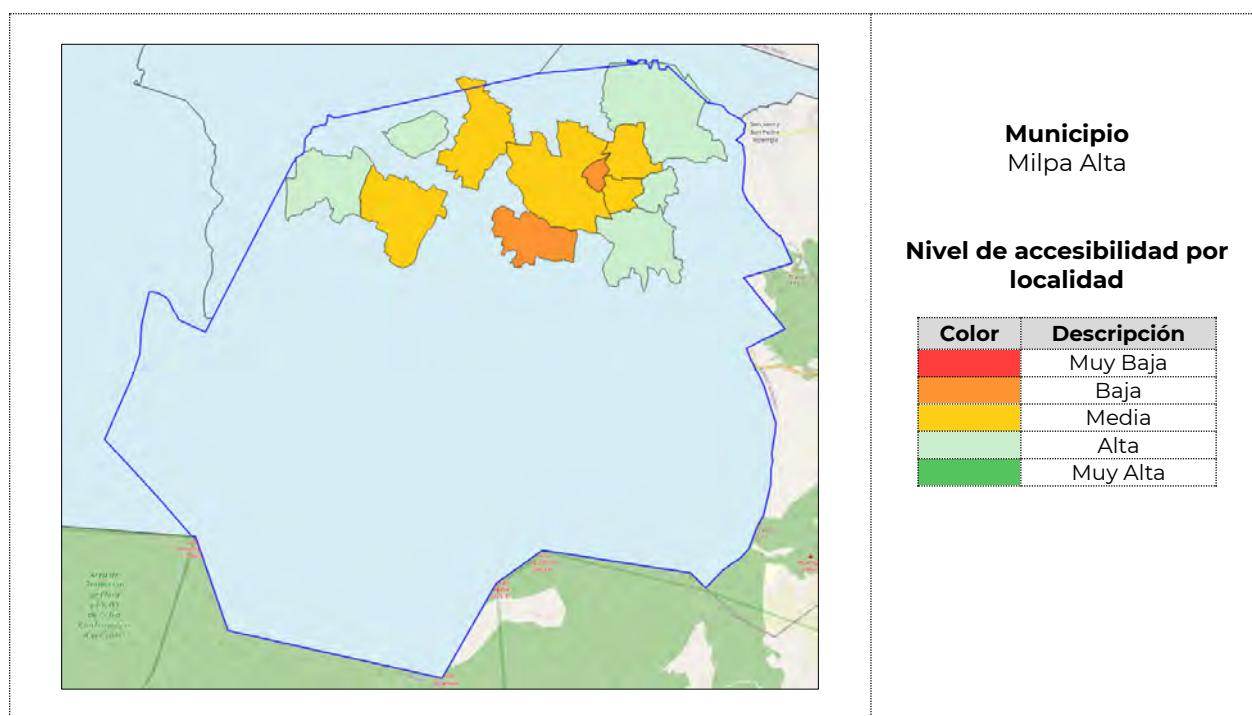


Figura 5.37. Indicador de accesibilidad: municipio de Milpa Alta

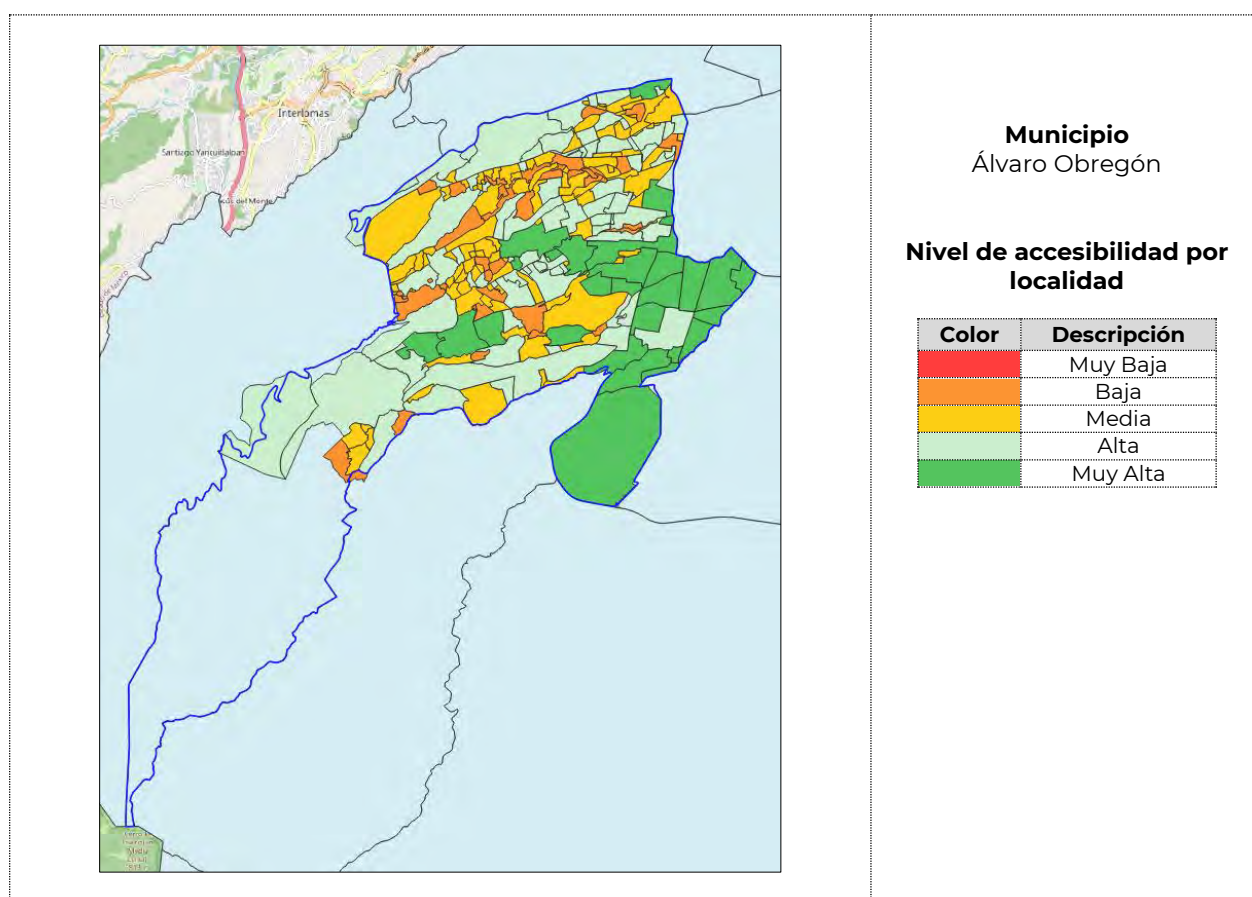


Figura 5.38. Indicador de accesibilidad: municipio de Álvaro Obregón

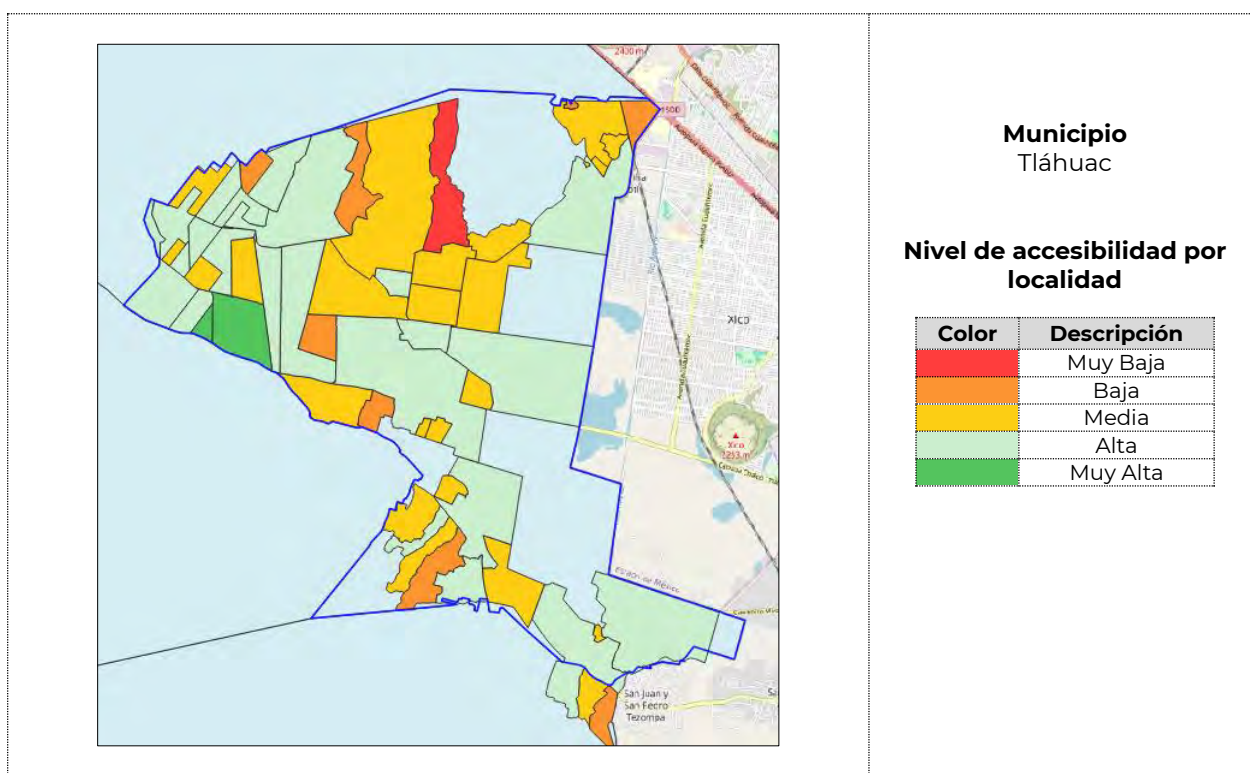


Figura 5.39. Indicador de accesibilidad por localidad: municipio de Tláhuac

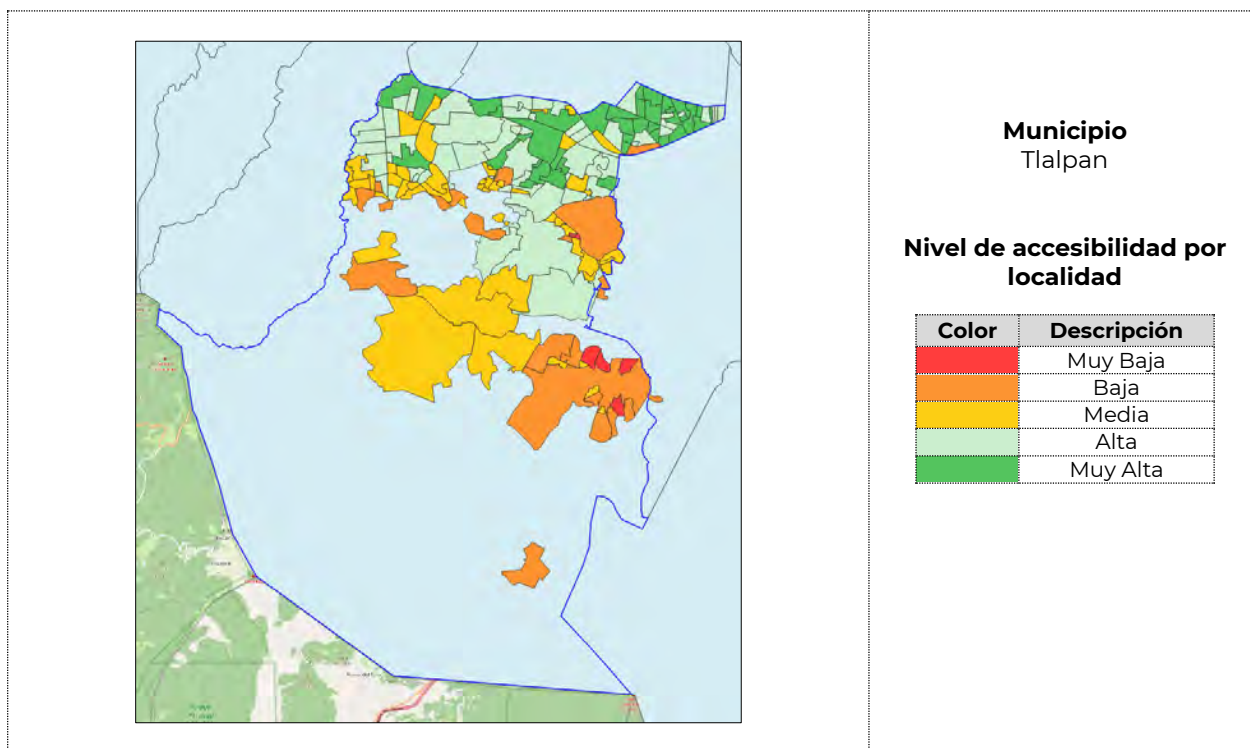


Figura 5.40. Indicador de accesibilidad: municipio de Tlalpan

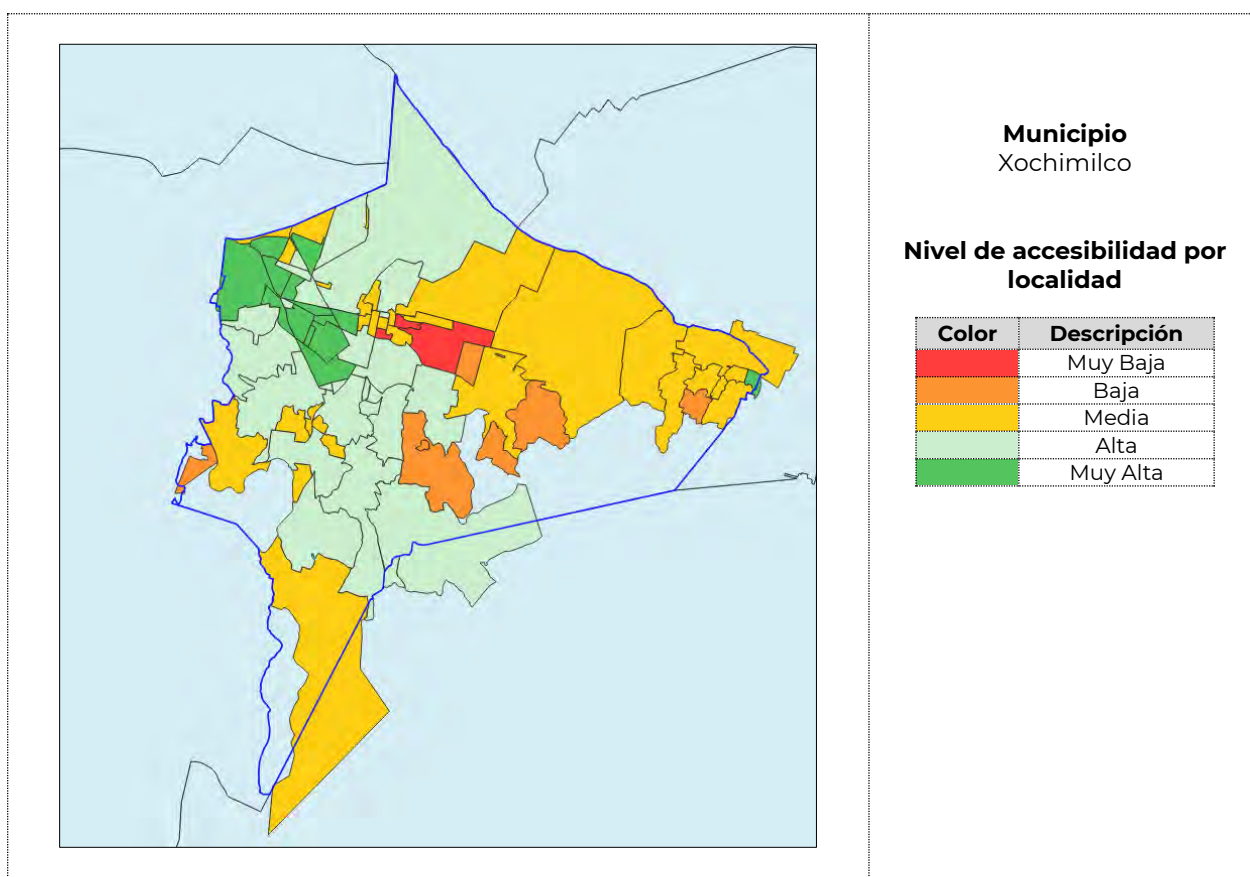


Figura 5.41. Indicador de accesibilidad: municipio de Xochimilco



Figura 5.42. Indicador de accesibilidad: municipio de Benito Juárez

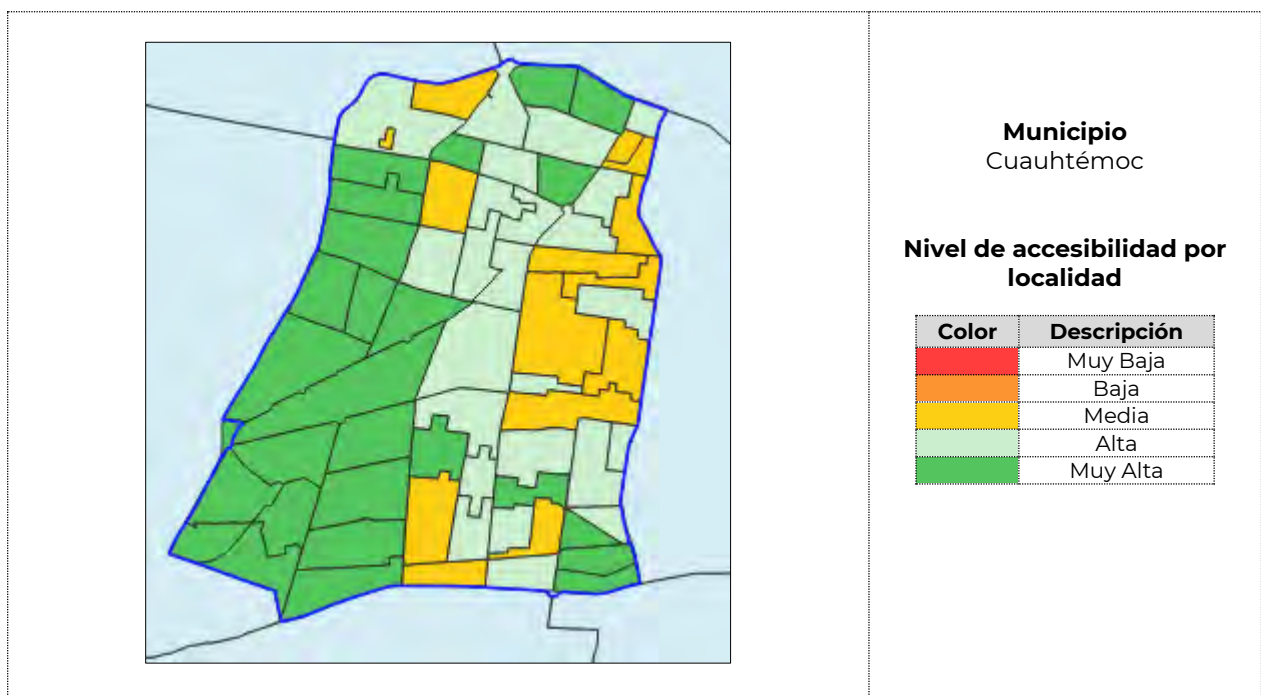


Figura 5.43. Indicador de accesibilidad: municipio de Cuauhtémoc

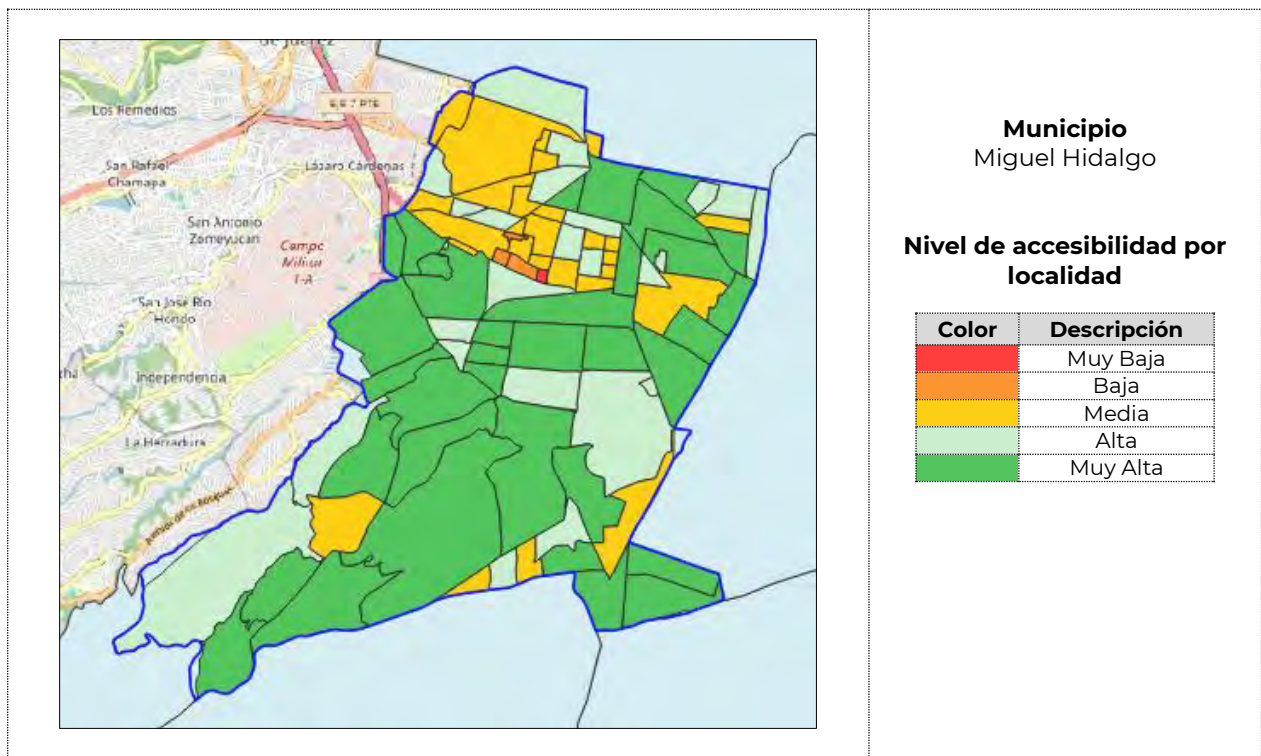


Figura 5.44. Indicador de accesibilidad: municipio de Miguel Hidalgo

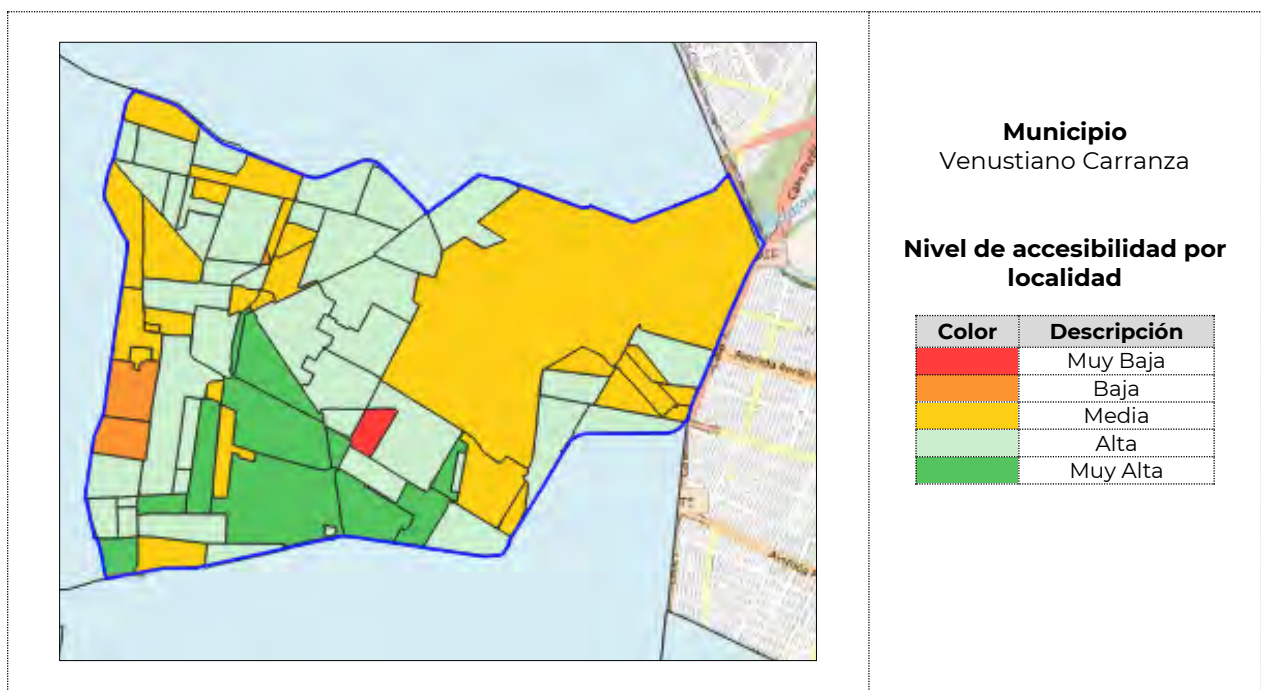


Figura 5.45. Indicador de accesibilidad: municipio de Venustiano Carranza

5.3. Conclusión

En este capítulo se ha llevado a cabo la validación del modelo matemático multicriterio, combinando distintos enfoques para garantizar su confiabilidad y aplicabilidad en la toma de decisiones. La validación incluyó un análisis comparativo con modelos existentes, lo que permitió evaluar la concordancia y consistencia de los resultados, además de identificar patrones de comportamiento en los datos. Para reforzar este proceso, se incorporó un análisis de sensibilidad con el objetivo de medir la estabilidad del modelo ante variaciones en los datos de entrada y en los valores de los pesos. Los resultados mostraron estabilidad ante datos que tienen el comportamiento similar a los datos recolectados; Sin embargo, para datos que tienen otros comportamientos el modelo muestra pequeños cambios. Este enfoque integral no solo permitió evaluar la robustez del modelo, sino también garantizar su capacidad de adaptación a diferentes escenarios y condiciones.

Por su parte, la interpretación del modelo ha sido un elemento fundamental, ya que no solo permitió estructurar y cuantificar la información, sino que también facilitó la identificación de patrones, la comparación y sobre todo la selección de alternativas. Esto proporciona una visión más clara del nivel de accesibilidad a los servicios y genera información valiosa para respaldar la toma de decisiones. En conjunto, estos enfoques integrados refuerzan la solidez del modelo y su utilidad para abordar problemas de accesibilidad desde una perspectiva multicriterio.

Capítulo 6. Conclusiones y trabajos futuros

En este punto se presentan las conclusiones obtenidas de la investigación, la cual se enfocó en el en el ámbito del análisis de accesibilidad a servicios educativos mediante un enfoque multicriterio.

6.1. Conclusiones generales

La presente investigación desarrolló un enfoque de Toma de Decisiones Multicriterio; orientada en accesibilidad a servicios, con énfasis particular en el acceso a escuelas primarias. Los principales hallazgos se sintetizan a continuación:

El Indicador de Accesibilidad a Servicios (IAS) es un modelo matemático multicriterio que permite una evaluación integral de las condiciones de accesibilidad a servicios educativos, al considerar simultáneamente un conjunto diverso de variables. Su diseño multidimensional agrupa dichas variables en tres criterios fundamentales: vivienda, educación y componentes espaciales, lo que evita un análisis parcial y permite captar distintas dimensiones del fenómeno. Cada criterio aporta información complementaria y relevante, favoreciendo así una interpretación más robusta, holística y objetiva de la accesibilidad territorial.

El método de ponderación de variables PCA-Entropía propuesto en esta investigación permitió asignar pesos de forma objetiva y precisa, reduciendo significativamente la dependencia de juicios subjetivos por parte de expertos. Además de garantizar una mayor objetividad, el método contribuye a incrementar la robustez y credibilidad del modelo multicriterio, al minimizar los efectos adversos derivados de la normalización o transformación de los datos, una limitante común en otros enfoques tradicionales. Asimismo, favorece la reducción de redundancia en la asignación de pesos, evitando sobrevaloraciones y promoviendo una representación más ajustada de la información disponible.

El enfoque Toma de Decisiones Multicriterio propuesto promueve un alto nivel de transparencia en los resultados del análisis. Lo anterior se logra mediante la implementación de un proceso estructurado; donde cada una de las etapas se encuentran documentada y pueden ser replicables. Esta característica convierte al modelo en una herramienta de caja blanca; favoreciendo su comprensión y análisis crítico por parte de tomadores de decisiones, ciudadanía y comunidad científica.

La aplicación integrada de distintas técnicas de selección de variables; donde a través de distintas perspectivas, permitió eliminar variables redundantes, identificar aquellas con mayor relevancia e interpretar de manera más clara las relaciones existentes entre ellas. La combinación de estos enfoques aportó elementos fundamentales para optimizar la precisión y consistencia de los resultados; asegurando que las variables seleccionadas fueran verdaderamente representativas y pertinentes dentro del análisis de accesibilidad.

La metodología desarrollada en esta investigación ha demostrado ser no solo adecuada para la evaluación y priorización de localidades en función de su accesibilidad a servicios, sino que también presenta un alto potencial de adaptación y aplicación en diversos contextos y niveles de análisis. Su estructura flexible permite

ampliar su utilización a distintas escalas territoriales, así como a la evaluación de otros tipos de servicios o infraestructuras. Esta capacidad de adaptación metodológica constituye un aporte relevante de este trabajo, al ofrecer una herramienta sólida, versátil y replicable que contribuye al fortalecimiento de los procesos de toma de decisiones en diferentes ámbitos.

Como resultado de esta investigación, se construyó un conjunto de datos especializado en accesibilidad a servicios, integrando variables espaciales, sociodemográficas y de vivienda, con el propósito de facilitar su uso en futuras investigaciones y análisis. Adicionalmente, se generó un conjunto de datos que incluye las variables consideradas en el modelo multicriterio, los valores de evaluación de las alternativas y la asignación final de las categorías obtenidas. Esta información también puede ser visualizada y analizada mediante representaciones cartográficas. Los archivos generados se encuentran disponibles en formatos shapefile (.shp) y CSV (.csv), y han sido alojados en un repositorio a través del siguiente enlace: <https://goo.su/fb0n6V4>.

6.1.1. Trabajos futuros

A partir de los resultados obtenidos en esta investigación, se identifican diversas líneas de trabajo futuro que permitirán ampliar y fortalecer los aportes del modelo propuesto:

Se contempla la implementación del Indicador de Accesibilidad a Servicios en una plataforma interactiva de consulta pública. Esta herramienta permitirá a tomadores de decisiones, ciudadanos y otros actores sociales visualizar de manera clara y dinámica los niveles de accesibilidad a servicios educativos, promoviendo así la transparencia, el análisis territorial y el diseño de políticas más informadas. A medida que se integren más datos o se desarrollen indicadores adicionales con enfoque territorial, será posible reforzar aún más la validación del modelo.

El modelo multicriterio enfocado en la accesibilidad a servicios constituye una base sólida para el desarrollo de modelos orientados a la optimización de recursos y la simulación de escenarios, particularmente en el contexto de la mejora de la accesibilidad educativa. En este sentido, Por ejemplo:

- o La implementación de modelos de optimización matemática para apoyar la distribución y gestión eficiente de recursos educativos.
- o La identificación de ubicaciones estratégicas destinadas a la construcción de nuevas infraestructuras escolares, con el fin de garantizar una adecuada cobertura en zonas de alta demanda.
- o La proyección de escenarios de demanda educativa futura mediante simulaciones, permitiendo anticipar las necesidades de infraestructura y planificar adecuadamente el crecimiento de la oferta educativa.

Se busca implementar el método de ponderación de variables propuesto en entornos computacionales (MATLAB, Python, R, etc.); utilizando la estructura matemática y el diagrama de proceso desarrollados como base para su codificación en diversos lenguajes de programación. Esta implementación permitirá evaluar su rendimiento

en diferentes conjuntos de datos y contextos, así como en aplicaciones a otros campos de investigación.

6.1.2. Logros obtenidos

Artículos JCR

- Pliego-Martínez, O.; Martínez-Rebollar, A.; Estrada-Esquivel, H.; de la Cruz-Nicolás, E. (2024). An Integrated Attribute-Weighting Method Based on PCA and Entropy: Case of Study Marginalized Areas in a City. *Applied Science*. 14, 2024. <https://doi.org/10.3390/app14052016>.
- De la Cruz Nicolás, E., Estrada-Esquivel, H., Martínez-Rebollar, A., Clemente-Torres, E. H., & Pliego-Martínez, O. A. (2025). Identification of variables influencing the origin of traffic congestion points. *Revista Facultad de Ingeniería Universidad de Antioquia*. <https://doi.org/10.17533/udea.redin.20250261>
- Colunga-Rodriguez, A. A., Martínez-Rebollar, A., Estrada-Esquivel, H., Clemente, E., & Pliego-Martínez, O. A. (2025). Developing a Dataset of Audio Features to Classify Emotions in Speech. *Computation*, 13(2), 39.
- De la Cruz-Nicolás, E., Estrada-Esquivel, H., Martínez-Rebollar, A., Pliego-Martínez, O. A., & Clemente, E. (2024). Index for Assessing the Performance Level of Vehicular Traffic on Urban Streets. *Urban Science*, 8(4), 204.
- De la Cruz-Nicolás, E., Esquivel, H. E., Rebollar, A. M., Pliego Martínez, O. A., & Hernández Pérez, M. Y. (2024). Representation of Traffic Congestion: A Multicriteria Analysis Approach Based on Distances. *Programming and Computer Software*, 50(8), 599-611.

Artículos de congreso

- Pliego-Martínez, O., Martínez-Rebollar, A., Estrada-Esquivel, H., De la Cruz-Nicolás, E., & Neftalí-Escobar Gómez, E. (2024). Comparative analysis using distance measures in GIS: Accessibility to hospitals. 3er Congreso Internacional de Tecnología y Ciencias Aplicadas (CITCA), Cuernavaca, Morelos, 20 de noviembre de 2024.
- Pliego-Martínez, O., Martínez-Rebollar, A., Estrada-Esquivel, H., & De la Cruz-Nicolás, E. (2024). Identificación de patrones en la selección de criterios mediante el análisis factorial: Marginación en localidades. Congreso Internacional de Investigación e Innovación Multidisciplinaria (CIIIM 2024), Orizaba, Veracruz.
- Pliego Martínez, O.; Martínez Rebollar, A.; Estrada Esquivel, H.; de la Cruz-Nicolás, E. (2024). Visualización de extensos conjuntos de datos espaciales con QGIS: Caso de estudio CDMX. *Tecnología y Ciencia Aplicada*. ISSN: 2992-8338.
- Nicolás, E. D. L. C., Esquivel, H. E., Rebollar, A. M., & Martínez, O. A. P. (2024). Impacto de la proximidad de los puntos de interés en la congestión de tráfico: Un enfoque

basado en la densidad de kernel. *Revista Digital de Tecnologías Informáticas y Sistemas*, 8(1), 88-97

- Ernesto De la Cruz-Nicolás, Hugo Estrada Esquivel, Alicia Martínez-Rebollar, Odette Alejandra Pliego Martínez. (2024). Identificación de variables causales de la congestión de tráfico mediante la prueba de Granger. Congreso Mexicano de Inteligencia Artificial.
- Ernesto De la Cruz-Nicolás, Hugo Estrada-Esquivel, Alicia Martínez-Rebollar, Odette Alejandra Pliego-Martínez, Rosa Olivia Maquinay-Díaz. (2024). Ponderación de variables causales de congestión de tráfico utilizando el método de entropía. 2º Congreso Internacional de Tecnología y Ciencia Aplicada 2024.
- De la Cruz-Nicolás, E., Estrada-Esquivel, H., Martínez-Rebollar, A., Pliego-Martínez, O. A., & Clemente, E. (2024). Methodology for the Development of Behavioral Models of Highly Dynamic Complex Systems. *International Journal of Advanced Computer Science and Applications*.
- Ernesto De la cruz-Nicolás, Odette Alejandra Pliego-Martínez, Jonathan Villanueva-Tavira, José Arnulfo Corona Calvario, Carlos Manuel Romano Acevedo. (2024). Recolección y almacenamiento de datos ambientales en la nube con ESP32 para Modelos Predictivos. 3er Congreso Internacional de Tecnología y Ciencias Aplicadas.
- De la Cruz-Nicolás, E., Estrada-Esquivel, H., Martínez-Rebollar, A., Pliego-Martínez, O. A. (2024). Modelado de velocidades vehiculares en función de la congestión de tráfico utilizando la Función Logística. Congreso Internacional de Investigación e Innovación Multidisciplinaria.
- De la Cruz-Nicolás, E., Martínez-Rebollar, A., Estrada-Esquivel, H., & Pliego-Martínez, O. A. (2023, November). Methodology to obtain traffic data and road incidents through maps applications. In *Ibero-American Congress of Smart Cities* (pp. 3-17). Cham: Springer Nature Switzerland.
- Pliego Martínez, O., Martínez Rebollar, A., Estrada Esquivel, H., y Flores Prieto, J. (2022). Midiendo la distancia de los servicios a través de la proximidad: caso de estudio CDMX. *AvaCient*, XII. No. 2. ISBN 2594-018X. Indizado por Latindex, Actualidad Iberoamerica, biblat, Fuente Académica Plus Database Coverage List, MIAR, LivRe, ESJI, Index Copernicus, entre otros.

Ponencias impartidas

- Análisis y visualización de datos con Sistemas de Información Geográfica. Instituto Tecnológico de Cuautla. 04 de abril de 2024.
- Las matemáticas de la ingeniería en el aprendizaje automático. Marco de la Jornada de Sistemas Computacionales 2024, Instituto Tecnológico de Cuautla. 21 de noviembre de 2024.

-
- Desarrollo de un modelo matemático con enfoque en la Toma de Decisiones Multicriterio: Caso de estudio accesibilidad a servicios. International Doctoral Consortium at MICAI 2024, Tonantzintla, Puebla, 22 de octubre de 2024.
 - Las matemáticas de la ingeniería en el aprendizaje automático. Marco de la Jornada de Sistemas Computacionales 2024, Instituto Tecnológico de Cuautla. 21 de noviembre de 2024.
 - Uso de tecnologías para la visualización de grandes cantidades de datos. 2º Ciclo de Conferencias Disciplinarias: Día del Programador. Universidad Tecnológica del Estado de Hidalgo. Tizayuca, Hidalgo, 14 de septiembre de 2022.
 - Modelo computacional para la determinación del grado de estrés en el acceso a servicios básicos. Marco del Taller de Cómputo Evolutivo Aplicado a Modelos Computacionales Orientados a Internet de las Cosas y Ciudades Inteligentes. Centro Nacional de Investigación y Desarrollo Tecnológico, 27 de octubre de 2021.

Talleres impartidos

- Introducción a los problemas de Optimización. Marco del 3er Congreso Internacional de Tecnología y Ciencias Aplicadas (CITCA). Centro Nacional de Investigación y Desarrollo Tecnológico. 2024. Cuernavaca, Morelos 2024.
- Introducción a las Redes Neuronales con Scilab. Marco del FLISOL 2014. Universidad Tecnológica del Estado de Hidalgo. Tizayuca, Hidalgo 2024.

Referencias

- Abdelaziz, S., Gad, M. I., & El Tahan, A. H. M. H. (2020). Groundwater quality index based on PCA: Wadi El-Natron, Egypt. *Journal of African Earth Sciences*, 172(October 2019), 103964. <https://doi.org/10.1016/j.jafrearsci.2020.103964>
- Aguirre, J. P. (2017). *Movilidad urbana en México*.
- Ahuja, R., & Tiwari, G. (2021). Evolving term “accessibility” in spatial systems: Contextual evaluation of indicators. *Transport Policy*, 113(March), 4–11. <https://doi.org/10.1016/j.tranpol.2021.03.006>
- Aït-Sahalia, Y., & Xiu, D. (2019). Principal Component Analysis of High-Frequency Data. *Journal of the American Statistical Association*, 114(525), 287–303. <https://doi.org/10.1080/01621459.2017.1401542>
- Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, 05(07), 754–767. <https://doi.org/10.4236/ojs.2015.57075>
- Azmoodeh, M., Haghighi, F., & Motieyan, H. (2021). Proposing an integrated accessibility-based measure to evaluate spatial equity among different social classes. *Environment and Planning B: Urban Analytics and City Science*, 48(9), 2790–2807. <https://doi.org/10.1177/2399808321991543>
- Baird, B. F. (1978). *Introduction to Decision Analysis*. Duxbury Press.
- Barríos, J. E., Moncada, C. A., & Escobar, D. A. (2023). Geographic Accessibility Analysis Using the E2Sfca Model in Hospitals Located in Armenia, Quindío. *GeoJournal of Tourism and Geosites*, 47(2), 648–655. <https://doi.org/10.30892/gtg.47233-1065>
- Basso, F., Frez, J., Martínez, L., Pezoa, R., & Varas, M. (2020). Accessibility to opportunities based on public transport gps-monitored data: The case of Santiago, Chile. *Travel Behaviour and Society*, 21(May), 140–153. <https://doi.org/10.1016/j.tbs.2020.06.004>
- Bierman, J., Bonini, C. P., & Hausman, W. H. (1996). *Análisis cuantitativo para la toma de decisiones*. Addison Wesley.
- Biosca, S. A. O., & Espinosa, M. A. A. (2018). Diseño metodológico para estimar indicadores de accesibilidad en entornos periféricos de una zona metropolitana. *Estudios Demográficos y Urbanos*, 33(1), 111–147. <https://doi.org/10.24201/edu.v33i1.1740>
- Bruno, G., Cavola, M., Diglio, A., & Piccolo, C. (2022). Geographical accessibility to upper secondary education: an Italian regional case study. *Annals of Regional Science*, 69(2), 511–536. <https://doi.org/10.1007/s00168-022-01146-6>
- Bucherie, A., Hultquist, C., Adamo, S., Neely, C., Ayala, F., Bazo, J., & Kruczkiewicz, A. (2022). A comparison of social vulnerability indices specific to flooding in Ecuador: principal component analysis (PCA) and expert knowledge. *International Journal of Disaster Risk Reduction*, 73(October 2021), 102897. <https://doi.org/10.1016/j.ijdrr.2022.102897>
- Cardozo, M. (2011). *Teoría de decisiones: Marco Teórico-Conceptual*. Universiada Autónoma Metropolitana.
- Carro, R. (2014). *Investigación de operaciones en administración*. Universidad Nacional de Mar del Plata. <https://nulan.mdp.edu.ar/id/eprint/2180/1/carro.2014.pdf>
- Cartone, A., & Postiglione, P. (2021). Principal component analysis for geographical data: the role of spatial effects in the definition of composite indicators. *Spatial Economic Analysis*, 16(2), 126–147. <https://doi.org/10.1080/17421772.2020.1775876>

-
- Chen, P. (2021). Effects of the entropy weight on TOPSIS. *Expert Systems with Applications*, 168, 114186. <https://doi.org/10.1016/j.eswa.2020.114186>
- CONAPO; IMT; CentroGeo. (2022). *Análisis geoespacial de la accesibilidad a centros urbanos de las localidades de México*. Secretaría de Gobernación, Consejo Nacional de Población (CONAPO), Instituto Mexicano del Transporte. <https://www.gob.mx/conapo/documentos/analisis-geoespacial-de-la-accesibilidad-a-centros-urbanos-de-las-localidades-de-mexico>
- Cortés, A; Rojas, L. (2020). Transporte, movilidad y accesibilidad: campos y métodos emergentes para el análisis geográfico contemporáneo. *Revista Científica de Ciencias Sociales y Humanas*, 79, 10–23. <https://doi.org/DOI:https://doi.org/10.33324/uv.vi79.433/>
- CURSOSONLINEGIS. (2025). *Mapas de coropletas: una guía para la clasificación de datos*. Análisis Multiespacial. <https://cursosonlinegis.com/mapas-de-coropletas-una-guia-para-la-clasificacion-de-datos/>
- Dejen, A., Soni, S., & Semaw, F. (2019). Spatial accessibility analysis of healthcare service centers in Gamo Gofa Zone, Ethiopia through Geospatial technique. *Remote Sensing Applications: Society and Environment*, 13(January), 466–473. <https://doi.org/10.1016/j.rsase.2019.01.004>
- Del, M. S. T. T., & Tabrizi, S. K. (2020). A methodological assessment of the importance of physical values in architectural conservation using Shannon entropy method. *Journal of Cultural Heritage*, 44, 135–151. <https://doi.org/10.1016/j.culher.2019.12.012>
- Eiselt, H. A.; Sandblom, C. L. (2022). Operations research, A Model-Based Approach. In *Communications on Pure and Applied Mathematics* (3rd ed., Vol. 8, Issue 1). Springer Nature Switzerland. <https://doi.org/10.1002/cpa.1955.3160080113>
- Eppen, G. D., Gould, F. J., Schmidt, C. P., Moore, J. H., & Weatherford, L. R. (2000). Investigación de operaciones en la ciencia de la administración. *Pearson*, 106–507–508.
- Erbaş, M., Kabak, M., Özceylan, E., & Çetinkaya, C. (2018). Optimal siting of electric vehicle charging stations: A GIS-based fuzzy Multi-Criteria Decision Analysis. *Energy*, 163, 1017–1031. <https://doi.org/10.1016/j.energy.2018.08.140>
- Faka, A. (2020). Assessing quality of life inequalities. A geographical approach. *ISPRS International Journal of Geo-Information*, 9(10). <https://doi.org/10.3390/ijgi9100600>
- Fuentes Bargues, J. L., Aragonés-Beltrán, P., Espinosa García, A., & Vivancos Bono, J. L. (2020). Multi-Criteria Evaluation of Alternatives for the Installation of Shw and Air Conditioning of a Health Centre in Valencia. *Proceedings from the International Congress on Project Management and Engineering, July*, 1476–1488.
- García, E.; García, H.; Cárdenas, L. E. (2013). Simulación y análisis de sistemas con ProModel. In *Sustainability (Switzerland)* (Vol. 11, Issue 1). Pearson. http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PE_MBTUNGAN_TERPUSAT_STRATEGI_MELESTARI
- García, J., Chagolla, H., & Noriega, S. (2006). Efectos de La colinealidad en el modelado de regresión y su solución. *Culcyt*, 16, 1–6. <http://erevistas.uacj.mx/ojs/index.php/culcyt/article/view/495/474>
- García Valdés, M., & Suárez Marín, M. (2013). El método Delphi para la consulta a expertos en la investigación científica Delphi method for the expert consultation in the scientific research. *Revista Cubana de Salud Pública*, 39(392), 253–267.

<http://scielo.sld.cu>

- Garrocho, C. F.; Campos, J. (2006). Un indicador de accesibilidad a unidades de servicios clave para ciudades mexicanas: fundamentos, diseño y aplicación. *Economía Sociedad y Territorio*. <https://doi.org/10.22136/est002006262>
- Garrocho, C. (1988). Análisis de la accesibilidad a los servicios de salud y de los sistemas de información geográfica: teoría y aplicación en el contexto del Estado de México. *Estudios Demográficos y Urbanos*, 8(2), 427. <https://doi.org/10.24201/edu.v8i2.883>
- Geoapify. (2024). *Geoapify location platform*. <https://www.geoapify.com/>
- Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, 12(2), 127–140. <https://doi.org/10.1016/j.jtrangeo.2003.10.005>
- Gewers, F. L., Ferreira, G. R., De Arruda, H. F., Silva, F. N., Comin, C. H., Amancio, D. R., & Costa, L. D. F. (2021). Principal component analysis: A natural approach to data exploration. *ACM Computing Surveys*, 54(4). <https://doi.org/10.1145/3447755>
- Gigović, L., Pamučar, D., Božanić, D., & Ljubojević, S. (2017). Application of the GIS-DANP-MABAC multi-criteria model for selecting the location of wind farms: A case study of Vojvodina, Serbia. *Renewable Energy*, 103, 501–521. <https://doi.org/10.1016/j.renene.2016.11.057>
- Gironés Roig, J., Casas Roma, J., Minguillón Alfonso, J., & Caihuelas Quiles, R. (2017). Minería de datos: Modelos y Algoritmos. In *Editorial UOC* (Issue 1). https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part%0Ahttp://www.editorialuoc.com
- González, M., Del Río, C., & Domínguez, J. M. (1989). Los servicios: concepto, clasificación y problemas de medición. *Ekonomiaz: Revista Vasca de Economía*, 13, 10–19. <https://dialnet.unirioja.es/servlet/articulo?codigo=1317365>
- Gray, S., & Cox, L. (2013). *Mental Modeler: A tool for environmental planning and research*. [http://www.mentalmodeler.org/articles/Mental Modeler Manual for Workshop.pdf](http://www.mentalmodeler.org/articles/Mental%20Modeler%20Manual%20for%20Workshop.pdf)
- Han, J., Wan, N., Brewer, S. C., & McCrum, M. (2023). Spatial access to Emergency General Surgery (EGS) services and EGS bypass behaviours in California. *Annals of GIS*, 29(1), 75–85. <https://doi.org/10.1080/19475683.2023.2165546>
- Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Planning Association*, 25(2), 73–76. <https://doi.org/10.1080/01944365908978307>
- Hasan, B. M. S., & Abdulazeez, A. M. (2021). A Review of Principal Component Analysis Algorithm for Dimensionality Reduction. *Journal of Soft Computing and Data Mining*, 2(1), 20–30. <https://doi.org/10.30880/jscdm.2021.02.01.003>
- IBM. (2024). *¿Qué son los datos geoespaciales?* <https://blogs.upm.es/espacio-geotig/sig/analisisvisualizacion-datos/2024>
- Indriyantho, B. R., Purnomo, J., Purwanto, Ottele, M., Han, A. L., & Gan, B. S. (2024). Multicriteria Sensitivity Analysis for Numerical Model Validation of Experimental Data. *International Journal of Technology*, 15(6), 1644–1662. <https://doi.org/10.14716/ijtech.v15i6.7146>
- Ingram, D. R. (1971a). The Concept of Accessibility; A search for an operational form.pdf. *Regional Studies*, 5(2), 101–107.
- Ishizaka, A., & Nemery, P. (2013). Multi-Criteria Decision Analysis: Methods and Software. In *Multi-Criteria Decision Analysis: Methods and Software*. <https://doi.org/10.1002/9781118644898>
- Janzamin, M., Ge, R., Kossaifi, J., & Anandkumar, A. (2019). Spectral Learning on Matrices

-
- and Tensors Suggested Citation. In *Foundations and Trends in Machine Learning* (Vol. 12, Issue 6). <https://doi.org/10.1561/22000000057>.Majid
- Jiang, J., Liu, J., Kadziński, M., Liao, X., & Dong, J. (2025). *Integrating Response Time and Attention Duration in Bayesian Preference Learning for Multiple Criteria Decision Aiding*.
- Jurado, M. F.; Salazar, E. J.; Samper, M. E.; Rosés, R. E. (2023). Estimación estocástica de estados operativos con un día en adelante en sistemas activos de distribución. *ERAC*, 1–8.
- Kaplan, N., Burg, D., & Omer, I. (2020). The spatial organization of accessibility and functional hierarchy: The case of Israel. *Computers, Environment and Urban Systems*, 80(April), 101429. <https://doi.org/10.1016/j.compenvurbsys.2019.101429>
- Kelobonye, K., Zhou, H., McCarney, G., & Xia, J. (Cecilia). (2020). Measuring the accessibility and spatial equity of urban services under competition using the cumulative opportunities measure. *Journal of Transport Geography*, 85(April), 102706. <https://doi.org/10.1016/j.jtrangeo.2020.102706>
- Keshavarz-Ghorabae, M., Amiri, M., Zavadskas, E. K., Turskis, Z., & Antucheviciene, J. (2021). Determination of objective weights using a new method based on the removal effects of criteria (MEREC). *Symmetry*, 13(4), 1–20. <https://doi.org/10.3390/sym13040525>
- Kherif, F., & Latypova, A. (2019). Principal component analysis. *Machine Learning: Methods and Applications to Brain Disorders*, 1(C), 209–225. <https://doi.org/10.1016/B978-0-12-815739-8.00012-2>
- Kim, J., Kim, D. H., Lee, J., Cheon, Y., & Yoo, S. (2022). A scoping review of qualitative geographic information systems in studies addressing health issues. *Social Science and Medicine*, 314(April), 115472. <https://doi.org/10.1016/j.socscimed.2022.115472>
- Knox, P. L. (1978). The Intraurban Ecology of Primary Medical Care: Patterns of Accessibility and Their Policy Implications. *Environment and Planning A: Economy and Space*, 10(4), 415–435. <https://doi.org/10.1068/a100415>
- Lardier, D. T., Blackwell, M. A., Beene, D., & Lin, Y. (2023). Social Vulnerabilities and Spatial Access to Primary Healthcare through Car and Public Transportation System in the Albuquerque, NM, Metropolitan Area: Assessing Disparities through GIS and Multilevel Modeling. *Journal of Urban Health*, 100(1), 88–102. <https://doi.org/10.1007/s11524-022-00701-y>
- Leung, C. K., Chen, Y., Hoi, C. S. H., Shang, S., Wen, Y., & Cuzzocrea, A. (2020). Big Data Visualization and Visual Analytics of COVID-19 Data. *Proceedings of the International Conference on Information Visualisation, 2020-Septe(Iv)*, 415–420. <https://doi.org/10.1109/IV51561.2020.00073>
- Logan, T. M., Williams, T. G., Nisbet, A. J., Liberman, K. D., Zuo, C. T., & Guikema, S. D. (2019). Evaluating urban accessibility: leveraging open-source data and analytics to overcome existing limitations. *Environment and Planning B: Urban Analytics and City Science*, 46(5), 897–913. <https://doi.org/10.1177/2399808317736528>
- López, F., & Aguilar, A. G. (2004). Niveles de cobertura y accesibilidad de la infraestructura de los servicios de salud en la periferia metropolitana de la Ciudad de México. *Investigaciones Geográficas*, 53, 185–209. <https://doi.org/10.14350/rig.30223>
- Mahmoudi, M. R., Heydari, M. H., Qasem, S. N., Mosavi, A., & Band, S. S. (2021). Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries. *Alexandria Engineering Journal*, 60(1), 457–464.

-
- <https://doi.org/10.1016/j.aej.2020.09.013>
- Mahuve, E. F.; Tarimo, B. C. (2022). Integrating fuzzy set function into floating catchment area measures: a determination of spatial accessibility of service points. *Annals of GIS*, 28(3), 307–323. <https://doi.org/10.1080/19475683.2022.2026477>
- Malczewski, J. (2017). Multicriteria Analysis. In *Comprehensive Geographic Information Systems* (Vol. 3). Elsevier. <https://doi.org/10.1016/B978-0-12-409548-9.09698-6>
- Malczewski, J., & Rinner, C. (2015). Multicriteria Decision Analysis in Geographic Information Science. In *Analysis methods* (Issue Massam 1993). http://www.amazon.com/Multicriteria-Decision-Analysis-Geographic-Information/dp/3540747567/ref=sr_1_1?ie=UTF8&qid=1430864854&sr=8-1&keywords=Multicriteria+decision+analysis+in+geographic+information+science
- Mandeville, P. B. (2008). ¿Por qué se deben centrar las covariables en regresión lineal? *Ciencia UANL*, 11(3), 300–305. d:%5CMarzo_2004%5CBibliografia%5CCovariables en regresi?n lineal.pdf%5CnD:%5CMarzo_2004%5CBibliografia%5CEjemplo_Mandeville.sav
- Marwal, A., & Silva, E. A. (2023). Analysing Inequity in Accessibility to Services with Neighbourhood Location and Socio-Economic Characteristics in Delhi. *SSRN Electronic Journal*, 1. <https://doi.org/10.2139/ssrn.4544867>
- Mendenhall W., Beaver R., & Beaver B. (2010). Introducción a la probabilidad y estadística. In *Cengage Learning*.
- Michalus, J. C., Sarache, W. A., & Hernández, G. (2015). Metodo de expertos para la evaluación ex-ante de una solución organizativa. *Visión de Futuro*, 19(1), 1–17.
- Montes, E., Cerezo, Y., & Romero, A. (2020). Eficiencia y Equidad espacial como principios para la planificación territorial de los establecimientos educativos: análisis de la Educación Primaria al sureste del municipio Maracaibo, Venezuela. *Anuario de La División Geografía*, 14.
- Moreno, J. M.; Aguaron, J.; Cano, F.; Escobar, M. T. (1998). VALIDEZ, ROBUSTEZ Y ESTABILIDAD EN DECISIÓN MULTICRITERIO. ANÁLISIS DE SENSIBILIDAD EN EL PROCESO ANALÍTICO JERÁRQUICO. *R.Acad.Cien.Exact.Fis.Nat.*, 92, 387–397. <https://docta.ucm.es/rest/api/core/bitstreams/5fe9da70-ce0f-4cf7-93c5-d04a3e012748/content>
- Moroto, C., Alcaraz, J., Ginestar, C., & Segura, M. (2012). Investigacion Operativa en Administracion y Direccion de Empresas. In *Angewandte Chemie International Edition*, 6(11), 951–952.
- Moskowitz, H.; Wright, G. P. (1979). *Operations Researchs Techniques for Management*. Prentice-Hall.
- Muñoz, B.; Romana, M.; Ordoñez, J. (2016). *Análisis de Sensibilidad de una Metodología de Decisión Multicriterio desarrollada para la Selección de Tipologías de Estructuras de Contención en una Autovía Urbana*. 0–5. <https://doi.org/10.4995/cit2016.2016.3491>
- Narváez, L. M. (2021). Análisis espacial de evaluación multicriterio con la lógica fuzzy y sistemas de información geográfica aplicado a la vulnerabilidad global en el departamento Pocito (San Juan-Argentina). *Geográfica Digital*, 18(36), 54. <https://doi.org/10.30972/geo.18364889>
- Németh, B., Molnár, A., Bozóki, S., Wijaya, K., Inotai, A., Campbell, J. D., & Kaló, Z. (2019). Comparison of weighting methods used in multicriteria decision analysis frameworks in healthcare with focus on low-and middle-income countries. *Journal of Comparative Effectiveness Research*, 8(4), 195–204.

-
- <https://doi.org/10.2217/cer-2018-0102>
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, 41(5), 673–690. <https://doi.org/10.1007/s11135-006-9018-6>
- Odu, G. O. (2019). Weighting methods for multi-criteria decision making technique. *Journal of Applied Sciences and Environmental Management*, 23(8), 1449. <https://doi.org/10.4314/jasem.v23i8.7>
- Ouma, P., Macharia, P. M., Okiro, E., & Alegana, V. (2021). Methods of Measuring Spatial Accessibility to Health Care in Uganda. *Practicing Health Geography*, 77–90. <https://doi.org/10.1007/978-3-030-63471-1>
- Padilla, M. A.; Llamuca, J.L.; Calderon, P. R.; Villamarín, J. M. (2019). Modelos matemáticos y la calidad del servicio aplicados al transporte urbano en Riobamba. *Ciencias Digital*, 3(2.2), 146–160. <https://doi.org/https://doi.org/10.33262/cienciadigital.v3i2.2.466>
- Park, J. D., Yoo, J., & Shin, W. Y. (2025). Criteria-Aware Graph Filtering: Extremely Fast Yet Accurate Multi-Criteria Recommendation. *WWW 2025 - Proceedings of the ACM Web Conference*, 4482–4493. <https://doi.org/10.1145/3696410.3714799>
- Pedrycz, W., Ekel, P., & Parreiras, R. (2010). Fuzzy Multicriteria Decision-Making. In *Fuzzy Multicriteria Decision-Making*. <https://doi.org/10.1002/9780470974032>
- Penchansky, R., & Thomas, J. W. (1981). The concept of access: Definition and relationship to consumer satisfaction. *Medical Care*, 19(2), 127–140. <https://doi.org/10.1097/00005650-198102000-00001>
- Prawda, J. (1996). *Metodos y modelos de investigacion de operaciones. Vol. 2 Modelos estocásticos.pdf*. Limusa.
- Pulido, H.; De la Vara, R. (2008). *Análisis y diseños de experimentos*. McGraw-Hill.
- Quezada, C. R., Bascuñán, M. M., De La Fuente Contreras, H., Faulbaum, A. S., Saéz, F. A., Me-Lla, G. F., Fuentes, C. P., & Cruz, J. C. M. (2019). Accessibility to equipment according to mobility and modes of transport in an average city, Los Angeles, Chile. In *Anales de Geografia de la Universidad Complutense* (Vol. 39, Issue 1, pp. 177–200). Universidad Complutense Madrid. <https://doi.org/10.5209/aguc.64682>
- Rădulescu, C. Z., & Rădulescu, M. (2018). A Hybrid Method for Cloud Quality of Service Criteria Weighting. *AIRO Springer Series*, 1, 425–432. https://doi.org/10.1007/978-3-030-00473-6_45
- Rehman, A., Khan, A., Ali, M. A., Khan, M. U., Khan, S. U., & Ali, L. (2020). Performance Analysis of PCA, Sparse PCA, Kernel PCA and Incremental PCA Algorithms for Heart Failure Prediction. *2nd International Conference on Electrical, Communication and Computer Engineering, ICECCE 2020, June*, 1–5. <https://doi.org/10.1109/ICECCE49384.2020.9179199>
- Ribbecca, S. (2025). *Mapa coroplético*. Catálogo de Visualización de Datos. https://datavizcatalogue.com/ES/metodos/mapa_coropletico.html
- Rincón, L. A. (2001). *Investigación de operaciones para ingenierías y administración de empresas*. Universidad Nacional de Colombia.
- Rodriguez-Segura, D., & Kim, B. H. (2021). The last mile in school access: Mapping education deserts in developing countries. *Development Engineering*, 6, 100064. <https://doi.org/10.1016/j.deveng.2021.100064>
- Rodríguez Cruz, J. (2020). Acercamiento al criterio de expertos en la investigación. *Primer Congreso Virtual de Ciencias Básicas Biomédicas En Granma*.
- Rong, P., Zheng, Z., Kwan, M. P., & Qin, Y. (2020). Evaluation of the spatial equity of medical facilities based on improved potential model and map service API: A case study in Zhengzhou, China. *Applied Geography*, 119(March), 102192.

-
- <https://doi.org/10.1016/j.apgeog.2020.102192>
- Roy, A., & Kar, B. (2022). A multicriteria decision analysis framework to measure equitable healthcare access during COVID-19. *Journal of Transport and Health*, 24(March 2021), 101331. <https://doi.org/10.1016/j.jth.2022.101331>
- Rubinstein, E.; Baladi, A. (2019). *Esto es lo que la accesibilidad implica en el diseño de políticas urbanas*. Moviliblog. <https://blogs.iadb.org/transporte/es/esto-es-lo-que-la-accesibilidad-implica-en-el-diseno-de-politicas-urbanas/>
- Şahin, M. (2021). A comprehensive analysis of weighting and multicriteria methods in the context of sustainable energy. *International Journal of Environmental Science and Technology*, 18(6), 1591–1616. <https://doi.org/10.1007/s13762-020-02922-7>
- Salehi, A., & Izadikhah, M. (2014). A novel method to extend SAW for decision-making problems with interval data. *Decision Science Letters*, 3(2), 225–236. <https://doi.org/10.5267/j.dsl.2013.11.001>
- Salem, N., & Hussein, S. (2019). Data dimensional reduction and principal components analysis. *Procedia Computer Science*, 163, 292–299. <https://doi.org/10.1016/j.procs.2019.12.111>
- Saraswat, S. K., & Digalwar, A. K. (2021). Evaluation of energy alternatives for sustainable development of energy sector in India: An integrated Shannon's entropy fuzzy multi-criteria decision approach. *Renewable Energy*, 171, 58–74. <https://doi.org/10.1016/j.renene.2021.02.068>
- Sarría, F. A. (2006). *Sistemas de información geográfica*. <https://www.um.es/geograf/sigmur/temariohtml/index.html>
- Schuschny, A., & Soto, H. (2009). Guía metodológica Diseño de indicadores compuestos de desarrollo sostenible Andrés Schuschny. In CEPAL. <https://hdl.handle.net/11362/3661>
- Scientific, E., & Company, P. (1976). *London. Particular attention is paid to the sensitivity of the measurement of accessibility to the choice of attractor variables , and of parameter values in the deterrence function . Attention is also focussed on the sensitivity of the measure to the z. 5.*
- Shahid, R., Bertazzon, S., Knudtson, M. L., & Ghali, W. A. (2009). Comparison of distance measures in spatial analytical modeling for health service planning. *BMC Health Services Research*, 9, 1–14. <https://doi.org/10.1186/1472-6963-9-200>
- Shao, Y., & Luo, W. (2023). Enhanced Two-Step Virtual Catchment Area (E2SVCA) model to measure telehealth accessibility. *Computational Urban Science*, 3(1). <https://doi.org/10.1007/s43762-023-00092-z>
- Silvestre Ortiz, J. R. (2019). *La consulta a expertos, un importante recurso en la investigación*. 181–192. <https://doi.org/10.5821/ctv.8241>
- Starr, M. K. (1971). *Systems Management of Operations*. Prentice-Hall.
- Sui, J., Zhang, G., Lin, T., Hamm, N. A. S., Li, C., Wu, X., & Hu, K. (2023). Quantitative Evaluation of Spatial Accessibility of Various Urban Medical Services Based on Big Data of Outpatient Appointments. *International Journal of Environmental Research and Public Health*, 20(6). <https://doi.org/10.3390/ijerph20065050>
- Taha, H. (2004). Investigación de operaciones. In *Investigación de operaciones (7a.)*. Pearson. <https://doi.org/10.18800/9789972429217>
- Tao, Z., Cheng, Y., & Liu, J. (2020). Hierarchical two-step floating catchment area (2SFCA) method: Measuring the spatial accessibility to hierarchical healthcare facilities in Shenzhen, China. *International Journal for Equity in Health*, 19(1), 1–16. <https://doi.org/10.1186/s12939-020-01280-7>

-
- Tinsley, H. E.; Brown, S. D. (2000). Handbook of Applied Multivariate Statistics and Mathematical Modeling. In *Munro's Statistical Methods for Health Care Research: Sixth Edition*. Academic press. <https://doi.org/10.4324/9781315816685-14>
- Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, 7(5), 437–456. [https://doi.org/10.1016/0038-0121\(73\)90041-4](https://doi.org/10.1016/0038-0121(73)90041-4)
- Wang, C. N., Le, T. Q., Chang, K. H., & Dang, T. T. (2022). Measuring Road Transport Sustainability Using MCDM-Based Entropy Objective Weighting Method. *Symmetry*, 14(5), 1–19. <https://doi.org/10.3390/sym14051033>
- Wang, S., Wang, M., & Liu, Y. (2021). Access to urban parks: Comparing spatial accessibility measures using three GIS-based approaches. *Computers, Environment and Urban Systems*, 90(September), 101713. <https://doi.org/10.1016/j.compenvurbsys.2021.101713>
- Wu, Y., Xia, L., He, J., Meng, Q., Yang, D., Liu, F., & Liu, N. (2023). Investigating Accessibility of Hospitals in Cold Regions: A Case Study of Harbin in China. *Health Environments Research and Design Journal*, 16(1), 142–160. <https://doi.org/10.1177/19375867221120201>
- Xia, N., Cheng, L., Chen, S., Wei, X. Y., Zong, W. W., & Li, M. C. (2018). Accessibility based on Gravity-Radiation model and Google Maps API: A case study in Australia. *Journal of Transport Geography*, 72, 178–190. <https://doi.org/10.1016/j.jtrangeo.2018.09.009>
- Yenisetty, P. T., & Bahadure, P. (2021). Assessing accessibility to ASFs from bus stops using distance measures: Case of two Indian cities. *Land Use Policy*, 108(May), 105567. <https://doi.org/10.1016/j.landusepol.2021.105567>
- Zafri, N. M., Sameen, I., Jahangir, A., Tabassum, N., & Hasan, M. M. U. (2021). A multi-criteria decision-making approach for quantification of accessibility to market facilities in rural areas: an application in Bangladesh. *GeoJournal*, 86(4), 1863–1879. <https://doi.org/10.1007/s10708-020-10161-z>
- Zardari; Noorul Hasan, Kamal, A., Sharif Monirussaman, S., & Yusop bin; Zulkifli. (2015). Weighting Methods and their Effects on Multi Criteria Decision Making Model Outcomes in Water Resources Management. In *Springer*.
- Zarghami; Szidarovszky. (2011). *Multicriteria analysis: applications to water and environment management*. Springer Science & Business Media. <https://doi.org/https://doi.org/10.1007/978-3-642-17937-2>